Detection mechanism of superconducting nanowire single-photon detectors

Engel, A ; Renema, J J ; Il'in, K ; Semenov, A

DOI: https://doi.org/10.1088/0953-2048/28/11/114003

Detection mechanism of superconducting nanowire single-photon detectors

A Engel[1], J J Renema[2], K Il'in[3] and A Semenov[4]

[1] Physik-Institut, University of Zürich, Winterthurerstr. 190, 8057 Zürich, Switzerland

E-mail: andreas.engel@physik.uzh.ch

[2] Leiden Institute of Physics, Leiden University, Niels Bohrweg 2, 2333 CA Leiden, the Netherlands

[3] Institute of Micro- and Nanoelectronic Systems (IMS), Karlsruhe Institute of Technology (KIT), Hertzstr. 16, 76187 Karlsruhe, Germany

[4] Institute of Optical Systems, German Aerospace Center (DLR), Rutherfordstrasse 2, 12489 Berlin, Germany

Abstract. In this paper we intend to give a comprehensive description of the current understanding of the detection mechanism in superconducting nanowire single-photon detectors. We will review key experimental results related to the detection mechanism, e.g. the variations of the detection probability as a function of bias current, temperature or magnetic field. Commonly used detection models will be introduced and we will analyze their predictions in view of the experimental observations. Although none of the proposed detection models is able to describe all experimental data, it is becoming increasingly clear that vortices are essential for the formation of the initial normal-conducting domain that triggers a detection event.

PACS numbers: 74.78.Na, 74.25.Wx, 85.25.Am, 85.25.Oj, 85.25.Pb

Keywords: single-photon detector, superconducting nanowire, vortex, vortex-antivortex pair, detection mechanism

Submitted to: *Supercond. Sci. Technol.*

1. Introduction

Superconducting nanowire single-photon detectors (SNSPDs or SSPDs) have attracted a lot of interest since they were introduced [1] and this interest is still growing. Their unique combination of speed, both in terms of high count rates and low timing jitter, high detection efficiencies, and low dark count rates makes them detectors of choice for a wide variety of very demanding applications as discussed in other articles in this Special Issue, see also [2–4].

On the other hand, these detectors are also interesting from a more fundamental point of view. Geometrically, the superconducting nanowire is for most of the H–T phase diagram best described as a two-dimensional (2D) system. The thickness d is typically about the same as the superconducting Ginzburg-Landau coherence length ξ_{GL}, whereas the width w is most often a factor of 10 or more larger. Only very close to the critical temperature T_c or for a few devices that were fabricated with extremely narrow strips does the width become comparable to the coherence length, and one might see the emergence of one-dimensional (1D) effects [5, 6]. In recent years many experimental studies have investigated the physics of the detection process in SNSPDs and several theoretical models have been proposed to describe the microscopic detection mechanism. Although currently no single model can reproduce all experimental results, the goal of this review will be to give a comprehensive summary of the current understanding of the detection mechanism in these detectors.

A detailed understanding of the detection mechanism is not only of fundamental interest. On a phenomenological level it is known that a reduction of the cross-section of the nanowire increases the sensitivity of the detector for low-energy photons, but at the same time results in a higher dark count rate. In recent years, a growing number of superconducting materials have been suggested with differing material parameters such as critical temperature or electronic density of states. Depending on the detection mechanism, these and other material parameters can have a profound influence on the detector performance. Without a sound understanding of the detection mechanism and how the material parameters influence the detector properties, it is a tedious task to find the best possible superconducting material for a specific application. A better understanding of microscopic processes involved in the detection of a photon may also aid in finding even better suited superconductors. Moreover, a fundamental understanding of the detection mechanism will make it possible to determine the ultimate limits of SNSPD performance, and thereby the feasibility of advanced applications.

It is becoming more and more evident that one has to include magnetic vortices in order to understand at least certain aspects of SNSPD, potentially they are even essential. In extended 2D superconductors magnetic vortices are the basic topological excitations and they have been studied in great detail and will not be discussed further, since a decent coverage of this topic would go way beyond the current work. It is important to remember though that vortices and anti-vortices (vortices with opposite orientation of the associated magnetic field) are present even in zero magnetic field, particularly in dirty, strongly type-II superconductors [7, 8], to which the superconducting materials belong that are typically used for SNSPD. A vortex-free state may again be realized in narrow superconducting strips or bridges with $w \lesssim \Lambda = 2\lambda_L^2/d$, where Λ is the Pearl length or effective magnetic penetration depth in films with $d \ll \lambda_L$ and λ_L the bulk London penetration depth [9], with some interesting consequences. The vortex-free Meissner state can be stable up to much higher fields than H_{c1} in bulk superconductors [10–12] and the critical-current density becomes width-dependent [13, 14], for example.

In SNSPD we have a special situation. The geometry of the superconducting structures is characterized by $\xi_{\mathrm{GL}}, d \ll w \ll \Lambda$ and their length $L \gg w$. They are usually operated in zero or ambient magnetic fields, therefore one can expect that there are no static vortices present in the superconducting structures. On the other hand, they are operated at bias currents I_b close to the experimental critical current I_c. The high bias current is necessary to achieve single-photon sensitivity, but makes the detectors also sensitive to fluctuations and dynamic vortex excitations. The situation becomes even more complicated in real devices by the typical meander structure. The sharp corners and turn-arounds disrupt the homogeneous current flow in the straight sections. It is therefore no surprise that it took many years to come from a simple, phenomenological description of the physics in these devices to a better, more microscopic understanding of the processes that are responsible for their remarkable detector characteristics.

The complete single-photon detection process can be split into several consecutive stages. It starts with the absorption of the photon in the superconducting structure and is mainly determined by the optical properties of the superconducting material and possibly surrounding dielectric and metallic layers [15, 16]. The superconducting structure can, for example, be imbedded into an optical cavity to increase the probability for photon absorption in the superconducting structure for a particular wavelength range [4, 17–19]. The absorption of the photon results in one electron being excited into an unoccupied state in the conduction band ‡. This excitation will relax and result in a local disturbance of the superconducting equilibrium state and the formation of an initial normal-conducting domain. A better understanding of the processes leading to the formation of this normal-conducting domain is the main topic of this review. The stages following the formation of the initial normal-conducting domain are again well understood [20, 21] and will not be treated in detail. In short, an electro-thermal feedback mechanism lets the normal-conducting domain grow rapidly due to self-heating by the bias current. Eventually the high resistance of the normal-conducting domain results in a significant reduction of the bias current through the superconductor. Superconductivity can then be recovered, the bias current returns to its original value and the detector is ready to detect the next photon.

We can now give a precise definition of the problem we want to address: what is the microscopic mechanism that leads to the formation of the initial normal-conducting domain that is the trigger for the voltage signal that can be registered. Presented experimental data have been obtained with different geometries of the nanostructures, see figure 1. Besides the traditional meander structures of practical detectors, bridge geometries (with and without bends) and so-called nanodetectors or bow-tie structures have been used as well. We will generally ignore details of the relaxation process of the first excited electron, instead a stationary disturbance will be postulated or a simplified phenomenological description will be used. In order to achieve our objective—to give a detailed review of the detection mechanism in SNSPD—we need a good understanding

‡ In general, SNSPDs have been used to detect visible and near-infrared photons with photon energies less than the work function of the superconducting material.

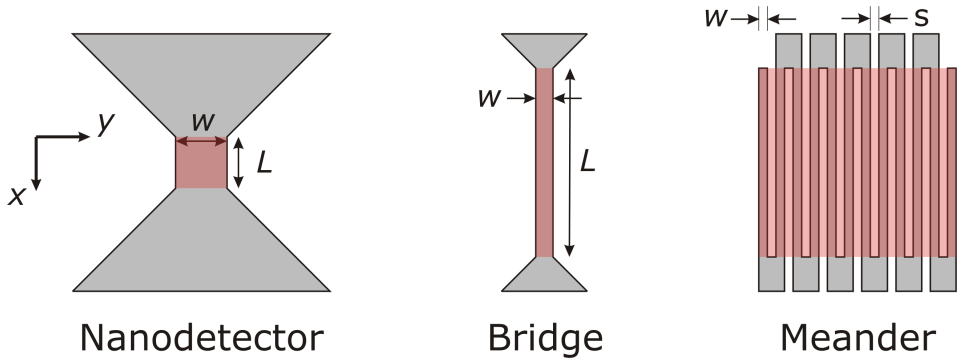

Figure 1. Schematic drawings of the three main structures, which have been used in experimental work (not to scale). Here we also introduce the coordinate system used in this review and give the definitions of the most important geometric parameters. For the nanodetector the width and length are comparable $L \lesssim w$, where as for the bridge and meander geometry we have $w \ll L$. The red shaded areas identify the active areas. The fill factor of the meander is given by $FF \approx w/(w+s)$.

of the stationary state of the superconducting nanowire before the photon absorption. This will be developed in the next section 2 followed in section 3 by a review of our understanding of dark counts in SNSPD. The main part of this paper will be sections 4 and 5. First we will give a review of the most important experimental observations that any detection model needs to be able to explain. This is followed by a presentation of proposed detection models and a thorough analysis of the models' results in view of the experimental data. We will conclude with a summary and a discussion of open issues.

2. Stationary state of SNSPD

The detection efficiency of an SNSPD grows when the operation current I_b increases [22]. A higher bias current corresponds to a larger cut-off wavelength or a lower minimum photon energy $E_{\min}$ [23] and also a smaller jitter [24]. The limit for I_b is an experimentally measured I_c which in turn cannot exceed the depairing critical current $I_{c,\mathrm{dep}}$. It has been shown experimentally that $E_{\min}$ decreases with an increase of the ratio of $I_c/I_{c,\mathrm{dep}}$, which has been adjusted by a variation of the stoichiometry of thin NbN films [25]. Therefore, the knowledge of the nature of the experimentally measured critical current is essential for device optimization. Although the two-fluid type of the temperature dependence of the critical current was often observed in earlier experiments [26, 27], this fact alone does not justify that the depairing current has been realized. Even if the linewidth remains unchanged, any deviation from the straight-line geometry reduces the measurable critical current due to current crowding and a corresponding local increase in the current density near bends or curves. In order to fill an area larger than the optical wavelength, the nanowire in SNSPD is usually patterned as a meander with sharp 180° turns. As stated above, the thickness d and the width w of typical SNSPD nanowires satisfy the conditions $d \lesssim \xi_{\mathrm{GL}}$ and $w \ll \Lambda$. Therefore, the current density remains uniform over the wire cross section in the straight portions of

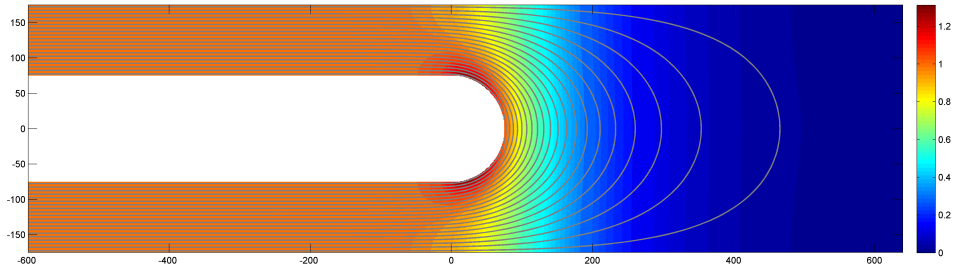

Figure 2. Computed current density (color coded and normalized to the homogeneous current density far from the turn) and stream lines (gray) in a 100 nm wide strip with a semicircular 180°-turn with radius $r = 75$ nm in the limit $\xi_{GL} \ll w$.

the meandering nanowire. The uniformity is disrupted at the turns, where the current crowds at the inner edges. The crowding increases the local current density above the mean current density in the straight portions of the wire, see figure 2. Therefore, the potential barrier for the vortex entry first disappears near the turns and this decreases the measurable critical current in comparison to the straight wire of the same cross section. Another consequence of the current crowding near turns is that the practically achievable ratio of the operation current to the depairing current in the straight portions remains less than it could be in a wire without turns. This makes straight portions less effective in detecting photons. Hence, decreasing the strength of current crowding near turns would greatly improve the spectral range of these detectors.

As predicted in [28], the current crowding effect in asymmetric superconducting structures with thickness $d \lesssim \xi_{GL}$ and width $w \ll \Lambda$ results in the suppression of the measured critical-current I_c with respect to the depairing critical current $I_{c,dep}$, which is characterized by the current reduction factor $R = I_c/I_{c,dep}$. The strength of the suppression is dependent on the particular geometry of the superconducting structure as well as on the angle and the radius r of the bend and the ratio of the coherence length to the width of the strip ξ_{GL}/w. Several experiments by different groups [29–31] have confirmed the predictions of [28]. It has been shown that the current crowding is stronger for wider superconducting strips but becomes weaker with an increase of the radius of the bend.

In figure 3(a) it is seen that j_c of 500 nm wide strips, which were made from a 4.5 nm thick TaN film on sapphire substrate, measured at zero magnetic field and $T = 4.2$ K decreases with an increase of bending angle α. This decrease of j_c is in good qualitative agreement with theoretical predictions for strips with $\alpha \leq 90°$ [28]. A further increase of the angle to 180° did not result in a further decrease of j_c as expected. The critical current densities of 90° and 180° bent bridges at zero magnetic field were almost the same, see figure 3(a). The latter phenomenon is due to the fact that in real structures the bend is not sharp § which leads to the weakening of the current crowding at the bend, once again, in qualitative agreement with the theory. The independence of the

§ The radius is usually larger than ξ_{GL}, determined by the resolution of the patterning technology.

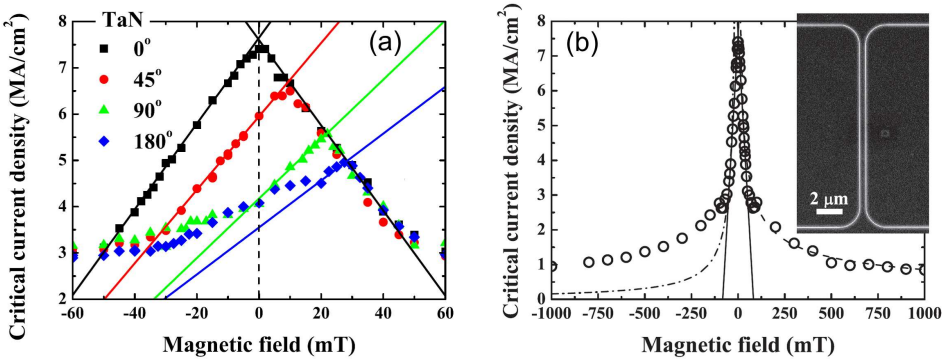

Figure 3. (a) Dependence of the critical current density on an applied magnetic field of 540 nm wide bridges of a 4.5 nm thick TaN film. Angles of the bends are indicated in the legend. Solid lines are least-square fits of vortex and anti-vortex branches of the curves to (1) and (2). (b) Dependence of the critical-current density of the TaN bridge without bend, 0° in (a), on a magnetic field applied perpendicular to the film surface over a much larger range compared to (a). Solid lines are the least-square fits of (1) to the experimental data. The dashed-dotted line shows the $j_c(B)$-dependence at $B > |B_{\rm stop}|$ calculated as $j_c(B) = 0.5 j_c(0) B_{\rm stop}/B$ as it is expected for a pure edge pinning mechanism. The dashed line is the best fit $\propto B^{0.5}$ to the experimental data at $B > |B_{\rm stop}|$ for the depinning critical current density as described in the main text. *Inset*: SEM image of a typical straight bridge. The white lines are reflections from the surface of the sapphire substrate. The black background is the superconducting film. (Reprinted figures with permission from [32]. © 2014 by the American Physical Society.)

critical current reduction factor R for angles $\alpha \geq 90°$ has been experimentally observed in strips of different superconductors (Nb, NbN, TaN) which were made from films with different thickness (see figure 1 in [32]).

There exists an optimal bending structure, which can be approximated by a semicircle with $r_{\rm opt} \approx w$, which should allow one to reach the critical current of the straight sections. However, an increase of the bending radius to the optimal values will lead to a decrease of the filling factor FF of a meander structure to $FF \leq 33\%$, which limits the absorption $ABS \propto FF$ [33] and thereby the detection efficiency of the detector $DE = IDE \times ABS$, with IDE being the intrinsic detection efficiency or detection probability and the assumption that no other parameter is changed. A reduction of the critical current with increasing FF has been observed experimentally [34] in line with the theoretical predictions [28]. This limitation may be overcome, if the particular application allows one to optimize the optical absorption for a narrow spectral bandwidth [35].

In contrast to the bias current, the Meissner currents, which are generated by externally applied magnetic fields, have opposite directions at opposite edges of a superconducting strip. Therefore, in straight strips subject to a bias current and an external magnetic field, the effective current density at one edge of the strip is higher than at the opposite edge. Thereby, the critical-current density corresponding to vortex penetration will be reached at smaller I_b independent of the direction of the magnetic

field, see figure 3(b). The $j_c(B)$-dependence is symmetric with a sharp maximum at $B = 0$ T and linear decrease for $B < B_{\mathrm{stop}}$. B_{stop} demarcates the Meissner (vortex free) and the mixed states of a superconducting strip [10] and evaluates in the London limit to $B_{\mathrm{stop}}(T) = \Phi_0/(2\pi e\xi_{\mathrm{GL}}(T)w)$ [36]. Therefore, it is dependent on the material (ξ_{GL}) and the geometry (w) of devices. In the Meissner state the critical current density is then

$$j_c(B) = j_c(0)\left(1 - \frac{B}{2B_{\mathrm{stop}}}\right), \tag{1}$$

which has been experimentally confirmed [32], solid lines in figure 3(b).

At magnetic fields $B > B_{\mathrm{stop}}$ the critical currents are no longer described by (1). The dash-dotted line describes the expected behaviour, if vortices are pinned only by the edge, given by $j_c(B) = 0.5j_c(0)B_{\mathrm{stop}}/B$ [36]. However, the experimental data are better described by the dashed line $\propto B^{0.5}$ as it may result from weak pinning in the superconducting film [37, 38].

However, in asymmetric structures the magnetic field can compensate the current crowding by the generation of a Meissner current with direction opposite to I_b at the inner radius of the bend. Of course, at the opposite edge of the strip the current density will be increased. However, the measured I_c will be increased up to a magnetic field B_{max}, for which the current densities at both edges become equal. Further increase of B above B_{max} will lead to a decrease of the measured I_c, compare to figure 3(a). This effect of a reduction of the current crowding effect by an externally applied magnetic field has been predicted theoretically [36] and confirmed experimentally in Al microstrips near T_c [39] and in Nb, NbN and TaN sub-micrometer bent strips at $T \ll T_c$ [32]. The increase of j_c with increasing B is not described by (1) anymore, but follows the dependence [36]

$$j_c(B) = j_c(0)R\left(1 - \frac{\delta}{R}\frac{B}{2B_{\mathrm{stop}}}\right), \tag{2}$$

where R is the already discussed factor of reduction of the critical current due to the current-crowding effect and δ is the field slope parameter (as it is defined in [36]). The latter factor describes a crowding effect for Meissner currents generated by an external magnetic field, therefore, δ is different from R. Both factors depend on the particular geometry of a sample; they are smaller than one and decrease with the increase of the bending angle. The critical current increases according to (2) until the intersection with the decreasing $j_c(B)$ according to (1) at the magnetic field B_{max}. It is remarkable that for $B > B_{\mathrm{max}}$ the $j_c(B)$-dependencies of bent bridges coincide with each other and follow the $j_c(B)$-curve of a straight bridge. At these magnetic fields vortices enter at the straight part of the bridge opposite to the inner corner of the bend and thus independent of the angle α [36].

3. Dark counts

Dark counts are a significant noise source of single-photon detectors. The noise-equivalent power is proportional to the square-root of the number of dark counts during

the sampling period. Therefore, high-sensitivity detectors need to have a dark count rate (DCR) as small as possible. Already very early publications on SNSPD demonstrated exceptionally low DCR compared with other single-photon detectors [40]. The most general definition of dark counts includes all events that are not caused by the absorption of a photon to be counted. This may include background photons from black body radiation or high energy quanta from cosmic rays, but also electronic noise due to unstable bias currents or electromagnetic interference. These extrinsic sources of dark counts have to be reduced as much as possible by appropriate filters and an optimized setup, for example, to a level that the DCR is given by intrinsic dark counts, only. From a practical point of view, the intrinsic DCR is the lower limit for a particular detector. On the other hand, it allows one to study the mechanisms that lead to intrinsic dark counts.

The first systematic measurements of DCR revealed two important characteristics of intrinsic dark counts: the DCR increases exponentially upon approaching the experimental critical current and decreases with decreasing operation temperature of the detector. This behaviour suggests that thermally activated fluctuations cause the formation of a resistive cross-section of the nanowire [41]. At sufficiently low bias currents one can often observe a cross-over to a constant or very slowly varying dark-count rate. Those dark-counts are usually attributed to electronic noise or a low level of background photons. Various mechanisms have been suggested to explain the observed exponential increase of the DCR: fluctuations of the order parameter [42, 43], thermally activated and quantum phase slips [44,45] and vortex excitations [46–48]. The thermally activated phase slips (TAPS) are the basic excitations of 1D superconducting wires and they can be successfully described by the LAMH-theory [49, 50], see [51] for an early review. It describes the temporary suppression of the superconducting order parameter in a minimal volume $A\xi_{\mathrm{GL}}$, A being the cross-section of the wire, which leads to a phase change of 2π, also known as Little's phase slips [52]. At high bias currents, phase slips and the resulting dark counts occur at a rate $\propto \exp\left(-\Delta F/k_B T\right) \exp\left(I_b/\tilde{I}\right)$, with k_B the Boltzman constant and $\tilde{I}$ a characteristic current scale [44, 51]. In this expression $\Delta F \propto H_c^2 A \xi_{\mathrm{GL}}$ is the saddle-point energy separating two meta-stable states, with H_c being the thermodynamic critical field. The thermally activated transition from one state to the other is accompanied by a phase slip of 2π, which is registered as a voltage pulse. Although this theory was developed for 1D wires with $A \lesssim \xi_{\mathrm{GL}}^2$, it can easily be extended to the case of thin superconducting strips by the appropriate expression for the cross-section $A = wd$. However, the saddle point energy for such phase slip lines (PSL) is proportional to the strip width, and that results in an exponentially decreasing rate of phase slips for increasing strip width.

As soon as the strip width exceeds about $4.4\,\xi_{\mathrm{GL}}$ magnetic vortices can exist inside the strip [53]. These can be single vortices due to an external magnetic field or the magnetic self-field generated by the bias current. They can also appear as vortex-antivortex pairs (VAP) as the topological excitations of a 2D superconducting

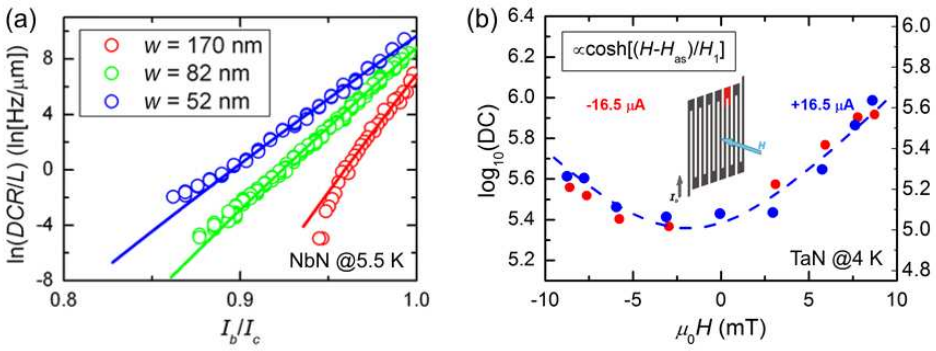

Figure 4. (a) Normalized dark-count rates for meanders of different strip width as indicated fabricated from a 6 nm thick NbN film. Fitting the analytical expression for VPS results in ξ_{GL} and Λ values consistent with parameters derived from resistivity measurements. Except for the widest structure for which the parameters differ by roughly a factor 1.5. Data are taken from [60]. (b) Magnetic field dependence of dark counts measured for a meander made from a 4.9 nm thick TaN film. Blue data points are results for one current direction, red data points for the reversed current direction. If one also reverses H-axes for the red data points, the two data sets coincide despite the obvious asymmetry with respect to the magnetic field. This behaviour can be understood and modelled assuming one turn-around with a reduced critical current, symbolized by the red section of the schematic inset. A detailed discussion can be found in [61].

film [54–57]. Such structures are comprised of closed magnetic flux loops and require a finite activation energy. Thus, at any non-zero temperature VAPs occur as thermodynamic fluctuations. These excitations are particularly frequent in strongly type-II superconducting thin films. In either case, one has to bear in mind that the vortices are distinct from free vortices inside an extended 2D film, because of the boundary conditions imposed by the edges of the strip. These boundary conditions result in an edge barrier that prevents vortex entry and/or exit in narrow superconducting strips [10–12, 58, 59]. It was argued in section 2 that the current for which the edge barrier for vortex entry vanishes limits the critical current. For lower bias currents vortices can overcome the finite edge barrier by thermal activation, and once they have overcome this barrier, the bias current exerts a Lorentz force that moves them across the strip. Such a vortex induced phase slip (VPS) can also be registered as a dark count event. A similar mechanism can lead to dark count events in the presence of VAP. Although the bias current does not exert a net force on a VAP, it results in a torque that orients the pair axis perpendicular to the current and lowers the binding energy [57]. Thermal activation can break up the pair and then the unbound vortex and antivortex move towards opposite edges of the strip and cause a dark count event [46].

Systematic theoretical and experimental investigations have found strong support for single vortices overcoming the edge barrier as the dominating mechanism leading to dark counts in SNSPD. In [62] numerical methods have been used to compare PSL and VPS. The key finding of this analysis is a critical width $w_c \approx (4.4 \pm 0.1)\xi_{\mathrm{GL}}$ that

coincides with the Likharev-criterion [53]. For strips with a width $w \leq w_c$ PSL are the only fluctuation mode leading to phase slips. For wider strips with $w > w_c$ phase slips caused by crossing vortices are also possible and with increasing w their free-energy barrier ΔF_{VPS} increases much slower than ΔF_{PSL}, particularly for the most relevant situation of high bias currents. The short coherence length in NbN implies a critical width $w_c \lesssim 20$ nm, such that in most devices studied to date, crossing vortices are expected to be far more frequent than PSL. A similar result, clearly favoring a vortex-based scenario over PSL, was obtained in an experimental study comparing dark count rates as a function of the reduced bias current of SNSPDs with different strip widths [60].

Analytical expressions have been derived for the three fluctuation modes for the specific case of SNSPD structures with $\xi_{\mathrm{GL}} \ll w \ll \Lambda$ and high applied bias currents [63]. These results suggest the VPS scenario as the most likely one to occur based on a comparison of the necessary activation energies to overcome the respective saddle-points. However, there is some controversy about the validity of this approach [64, 65]. First, the London-model was used for the description of vortices, which is known to fail when vortices are too close to the edge of the strip. Second, the energy associated with the vortex core has been neglected. Nevertheless, they fit their analytical formula for VPS to the data of [60] and obtain very consistent fit parameters except for the widest meander line with $w \approx 170$ nm, see also figure 4(a). As discussed in [62], there may be a cross-over from the VPS dominated fluctuations to VAP as the strip width is increased. The rate of single vortices entering the strip is proportional to the length of the wire, whereas the rate of VAP break-up is proportional to the area of the superconducting film. Therefore, the ratio of VAP to VPS events is proportional to the strip width w and one can expect that VAP events start to dominate the dark count events for strips wider than a given cross-over width. However, this has never been systematically investigated.

The problem of dark counts and which mechanism is responsible for the break-down of superconductivity in SNSPDs has also been investigated using the Ginzburg-Landau (GL) equation and a numerical relaxation method to find stationary states [65]. It is found that for strips typical for SNSPDs and $I_b/I_{c,\mathrm{dep}} \lesssim 0.6$ the results of the London-model are reproduced. At higher currents this approach comes to slightly different results, it even results in PSL requiring the lowest activation energy for bias currents very close to $I_{c,\mathrm{dep}}$. The current-crowding effect in the bends of the typical meander structure is also reproduced which favors dark counts to occur near bends.

If single vortices are responsible for the observed intrinsic fluctuations in SNSPD, one would expect a distinct magnetic-field dependence of the dark-count rate, whereas the dark-count rate caused by VAP should be independent of the magnetic field. In a follow-up publication Bulaevskii *et al.* extended their model to include the influence of an external magnetic field [66]. For small applied fields $H < H^*$, with H^* the vortex-entry field at $I = 0$, they predict the following simple relation:

$$R_{DC}(I, H) = 2R_{DC}(I, 0)\cosh\left(H/H_1\right), \tag{3}$$

where I and H are the applied current and magnetic field, respectively, and H_1 is the

characteristic field scale of a few mT for a typical SNSPD meander. This functional dependence of the dark-count rate R_{DC} has been confirmed in TaN and NbN SNSPD up to ~ 100 mT [61,67], as shown in figure 4(b). Furthermore, these studies have found evidence that the dark counts originate in the turnarounds of the meander structure. This is a consequence of the current-crowding in sharp bends as discussed in section 2, which leads to a reduced vortex-entry barrier at the inner bends of the turnarounds as compared to the energy barrier along the straight sections of the meander.

Thus, a consistent picture has emerged that single vortices which overcome the edge barrier are the dominating mechanism leading to dark-count events in SNSPD, at least at temperatures of around 4 K, for which most of the experimental investigations were done‖. Most experiments and theoretical models are consistent with dark counts originating near a weak point. Naturally, these are the bends in the typical meander structures of SNSPDs. However, some open questions and interesting observations remain to be clarified. A temperature-dependent study [69] revealed a growing inconsistency between theory and experiment at temperatures below ≈ 2.5 K¶. Very recently, the switching current distributions as a function of temperature in typical meander structures as used for SNSPD have been measured [70]. A detailed statistical analysis suggests that at temperatures $T \lesssim 2$ K dark counts may not be thermally activated, but caused by quantum-mechanical tunneling events, so called QPS. Quantum tunneling of vortices through the edge barrier was claimed previously [71] and certainly deserves further investigations. A possible dependence of the dark-count rate on the thermal coupling of SNSPDs to the thermal bath has been observed by Hofherr *et al.* [72]. It appears that the attempt rate depends on the thermal coupling and that it can be significantly reduced by better thermal contact of the SNSPD to its mechanical support and thermal bath. This effect cannot be understood within the current models.

4. Photon counts

What are the processes after the absorption of a photon that lead to a detection event in SNSPD is the key question, which we are going to address in the following. There exists a large amount of experimental observations that a good detection model must be able to explain. In order to set the stage, we will first review the most important results of "classical" experiments. Very distinct results have been obtained with quantum detector tomography (QDT), which we will discuss in a separate section, before we will give an account of recent experiments on the temperature, position, and magnetic-field dependence of photon counts in SNSPDs.

‖ This role of vortices for the generation of dark counts has been further strengthened in a very recent publication [68] that became available after the first submission of this paper, in which a significant reduction of the dark-count rate in ferromagnet-superconductor bilayers is reported.

¶ Contrary to the conclusions of the authors of [69], their data can be successfully fitted to the VPS model of [63], with very reasonable fitting parameters for $T \geq 2.5$ K. These fitting parameters become increasingly unrealistic at lower temperatures.

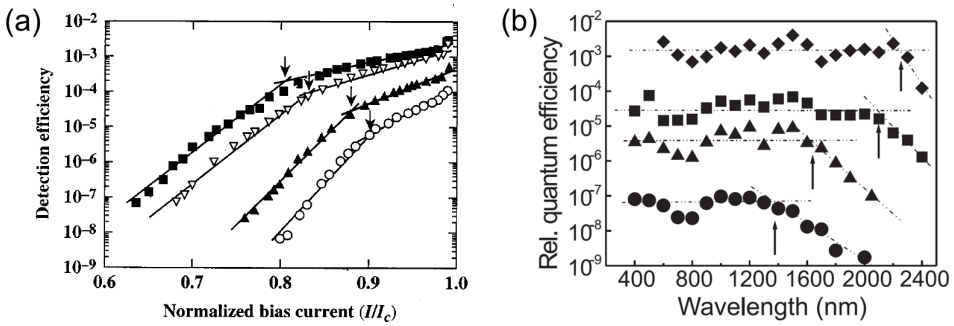

Figure 5. Two examples of early measurements of the count-rate dependence. In (a) the detection efficiency was measured as a function of reduced bias current for different photon wavelengths from 560 to 1260 nm (Reproduced with permission from [73]. © 2002, AIP Publishing LLC.) and in (b) the normalized detection efficiency vs. photon wavelength is shown for reduced bias currents I_b/I_c from 0.6 to 0.89 (Reproduced with permission from [74]. © 2005, Springer Verlag). In both graphs lines are drawn for the different detection regimes above and below the threshold and arrows mark the thresholds.

4.1. Detection threshold in SNSPDs

The typical characterization of an SNSPD is done in either of two ways. One measures the photon-count rate at a fixed photon wavelength and a varying bias current or vice versa with a fixed bias current while one changes the wavelengths of the photons, see examples in figure 5. These measurements are complementary to each other. The first one has the advantage that the absorptance is constant and that it is easier to control the photon rate and keep it constant, particularly when changing the temperature or when comparing different devices. It is possible to extract minimum photon-energy ($E_{\min}$) threshold-current ($I_{\rm th}$) pairs from these experiments. For higher bias currents and/or photon energies the photon-count rate is approximately constant (plateau) and below the threshold the photon-count rate is rapidly decreasing following an approximately exponential relation. The exact mathematical relation between $E_{\min}$ and $I_{\rm th}$ will be an important test of the detection models. The generally observed behaviour shows that one needs a lower $I_{\rm th}$ to detect a higher energy photon and vice versa.

When comparing the threshold current with theoretical models, at different temperatures or between different devices, the most convenient and practical quantity is not the absolute $I_{\rm th}$, but instead the reduced $I_{\rm th}/I_{c,\rm dep}$, which is scaled by the depairing critical current. It can lead to confusions and make comparisons difficult, if $I_{\rm th}$ is scaled by the experimental critical current. As we have seen in section 2, the experimental critical current I_c is generally device-dependent, given by the geometry and possibly microscopic defects or variations of the cross-sectional area of the strip, for example. Therefore, it is advisable to always use the depairing critical current $I_{c,\rm dep}$ as the most relevant current scale.

There is always some rounding near the threshold and the plateau may not be characterized by a truly constant count rate. Both of these non-ideal behaviours

may be at least partly explained by device inhomogeneities, such as small strip width and thickness variations [75] or locally varying superconducting gap [76]. However, the rounding near the threshold persists even for the most homogeneous devices, characterized by an almost perfectly constant count rate above the threshold. In the plateau region, it appears to be possible that the detection efficiency is only limited by the absorptance of photons in the thin superconducting film, i.e. every absorbed photon is detected, in other words $IDE = 1$ [77, 78].

The threshold itself depends on several material and device parameters. E_{min} and I_{th} are expected to increase with an increase in width or thickness of the superconducting meander [23, 77–80], based on the fact that the volume that is needed to switch into the normal-conducting state for triggering a detection event increases with increasing cross-section. For simplicity, we will call a change of the threshold to higher E_{min} or higher $I_{\mathrm{th}}/I_{c,\mathrm{dep}}$ an up-shift of the threshold and correspondingly a change in the opposite direction a down-shift. The investigation of the influence of material parameters on the detection threshold is experimentally more difficult. Generally, several possibly relevant parameters change for different materials such as T_c, the density of states $N(0)$, or the electron diffusion constant D. The dependence on T_c seems obvious. A lower T_c, and thus a lower superconducting gap energy Δ, results in a larger number of quasi-particle excitations. This should result in a down-shift of the threshold. Comparing the characteristics of a growing number of SNSPDs made from different materials confirms this expected trend [81–87].

A more quantitative comparison is very difficult, since the device parameters and experimental setups are not comparable, although in some studies [83, 84] a direct comparison between at least two materials was attempted. In an early experiment [74], Nb and NbN SNSPD were compared with each other. For that study the NbN film was deliberately fabricated with a non-stoichiometric composition to achieve a T_c comparable to the thin Nb film. This study revealed the importance of additional material parameters such as $N(0)$ and D. In order to achieve a low E_{min} threshold for a given bias current, available experimental data suggest to not only use a low-T_c material, but also a superconductor with a low $N(0)$ and a low D. This combination of material parameters results in a large number of excess quasi-particles (T_c), a low density of Cooper pairs and thus also a low density of quasi-particles required to suppress superconductivity (T_c and $N(0)$), and a small diffusion length that maximizes the density of quasi-particles near the absorption site (D).

The maximum number of excess quasi-particles and their local density as a function of time is also influenced by characteristic time scales of the superconducting material. These parameters are relatively well known for NbN and Nb [74, 88] but much less well known for the other materials, and no systematic studies of these time scales on the threshold have been published so far. Depending on the detection mechanism there may be additional influences on the threshold through the superconducting coherence length ξ_{GL} and magnetic penetration depth λ_L. However, these parameters are microscopically linked to T_c, $N(0)$, D and the electron mean free path l (see e.g. [60] and references

therein), therefore they cannot be varied independently.

Detailed information about detection threshold can be experimentally obtained using QDT. The technique and important results obtained with it will be discussed in the following section.

4.2. Linear energy-current relation from QDT

QDT [89, 90] is a powerful tool which has proven to be invaluable for the investigation of the detection mechanism in SNSPDs. It is an experimental procedure to measure the detection statistics of a photon detector whose response is unknown. The goal is to find the response of the detector in the number state (Fock) basis, i.e. to find the probability of a detection event p_n, if the detector were to receive exactly n photons as an input. A brief review of the technique, its foundation and prerequisites are given in the appendix. In the following we will discuss the experimental results in relation to SNSPDs.

In [91], a tomography protocol which compensates for external optical losses (see section Appendix A.4) was applied to a 150 nm wide NbN nanodetector [92], for a wavelength of 1500 nm. Figure 6(a) shows the results, which were obtained. The key result in this figure is the almost constant value of the linear efficiency η. This parameter varies by a factor 1.5, in the bias current range from 11 to 20 μA, while the count rate changes by many orders of magnitude. This confirms that the separation between p_n and η is robust. Moreover, the value of the linear efficiency is consistent with the optical cross-section of the device.

This leads to the following interpretation: the linear efficiency η contains the optical losses outside the detector (i.e. before photon absorption) whereas the nonlinear effects (which occur after photon absorption) are contained in the p_n. It should be noted that there can also be linear loss effects inside the detector. For the samples discussed here, they are not observed, but they have been seen in other samples [93].

In [94], the experiment discussed above was repeated for wavelengths of 1300 and 1000 nm. This data was combined with the original data from [91] at 1500 nm. The result of this is shown in figure 6(b).

In figure 7, the results presented in figure 6(b) are scaled by a linear relation

$$p_n(E, I_b) = p_n(I_b + \gamma E). \tag{4}$$

It should be noted that the results superimpose. This demonstrates that there is a *universal curve* for NbN detectors: once the detection probability as function of bias current for one wavelength is known, it can be predicted for other wavelengths by using the scaling relation. Furthermore, the curves superimpose regardless of the photon number involved in the excitation. This demonstrates that only the total amount of energy in the excitation matters, and not the photon number, i.e. the way in which the energy is distributed over several photons.

The scaling relation (4) suggests a linear relation between the photon-energy and the threshold current necessary to achieve a given detection probability $p_n \leq 1$. In order to confirm this observation and to increase the energy range, an experiment on a slightly

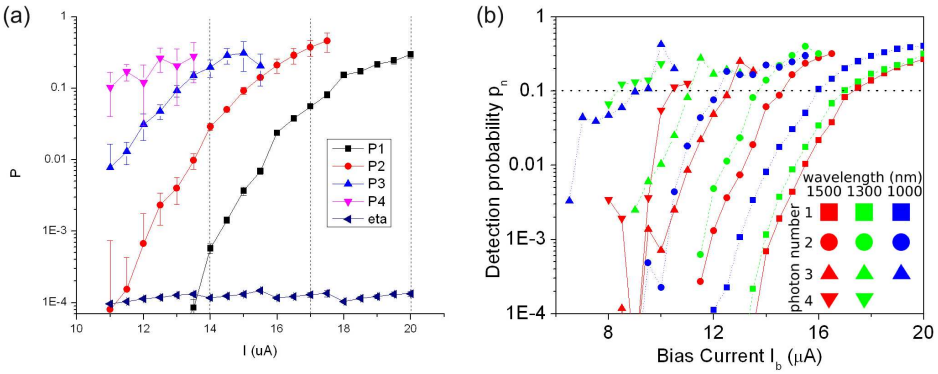

Figure 6. (a) Quantum detector tomography on a 150 nm wide nanodetector with a wavelength of 1500 nm. Both the linear efficiency and the nonlinear detection probabilities are shown (Reproduced with permission from J. J. Renema *et al.*, Opt. Express, 20, 2806, 2012.). (b) Quantum detector tomography on the same detector with the additional wavelengths of 1000 and 1300 nm. The detection probability is shown for up to 4 photons for 1300 nm and 1500 nm, and for up to 3 photons for 1000 nm (Reproduced with permission from [94]. © 2013, American Physical Society.).

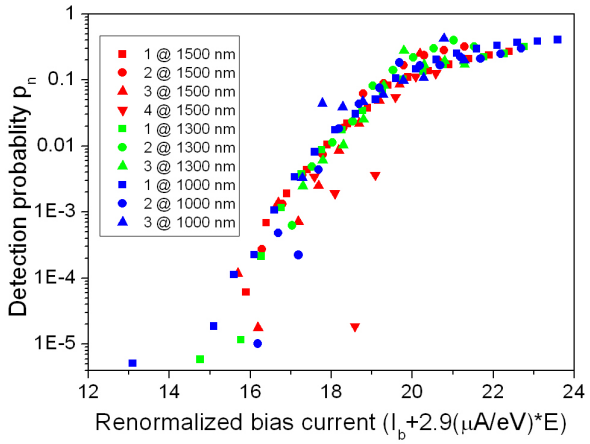

Figure 7. Demonstration of the scaling behaviour of the detection probability curves. The data of figure 6 have been scaled according to (4). They superimpose and describe the universal response curve for the nanodetector (Reproduced with permission from [94]. © 2013, American Physical Society.).

wider (220 nm) nanodetector was conducted, over a series of wavelengths from 460 nm to 1650 nm [95]. The results of this experiment, presented in figure 8, confirm that the relation between photon energy and bias current is indeed linear over at least one order of magnitude in photon energy.

Figure 8 shows all combinations of photon energy and threshold bias current I_{th} required to produce a detection probability of 1%. The arrows identify the points where excitations of different photon numbers with equal energy coincide, confirming the idea that only the overall energy of the excitation matters for the detection probability and

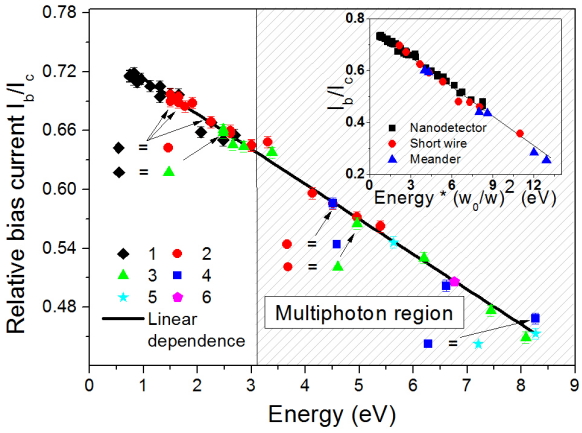

Figure 8. Linear dependence between bias current and photon energy. *Inset:* the energy-current relation of three different detectors with different geometries (Reproduced with permission from [95]. © 2014, American Physical Society.).

validating the use of multiphoton excitations to increase the accessible range of energies. The inset shows a comparison between three different samples: the nanodetector from the main figure, a short wire and a meander detector of the type used in applications. The results of all three samples are described by the same linear relation. Extrapolating the linear dependence

$$I_{\rm th} = I_0 - \gamma E \tag{5}$$

to $E = 0$, results in a value of $I_0 \approx 0.75 I_c$. It is important to note that the extrapolated I_0 depends weakly on the chosen value of p_n to determine $I_{\rm th}$.

There might be deviations from this linear relation for low photon energies E. Specifically one detection model, see section 5.5, predicts a strong nonlinearity as $E \to 0$. It remains an experimental challenge to extend the measurements to lower photon energies.

4.3. Temperature dependence of the detection threshold

It was realized early on that a reduction of the operation temperature of NbN SNSPD from $\sim$ 4 K to $\sim$ 2 K does not only reduce the dark-count rate, but also results in a down-shift of the threshold [41]. The downshift of the threshold continues to even lower temperatures well below 1 K [96]. This is a counter-intuitive observation. A lower operation temperature results in a larger gap energy Δ and thus a smaller number of excess quasi-particles, which should cause an up-shift of the threshold. Despite repeated observations of this temperature dependence of the detector performance, it was not until very recently that a systematic study was conducted that confirmed this trend at least down to 0.5 K [97]. Increasing the temperature above 4 K $\approx 0.5 T_c$ results in a fast deterioration of detector performance and increasing instability, at least for NbN and TaN detectors. It appears that WSi detectors can be operated at least up to $0.7 T_c$ [98].

In [95] the temperature dependence of the linear energy-current relation has been investigated using QDT. The authors have found a temperature-independent γ and an I_0 that is slowly decreasing with increasing T over a temperature range that spans from about 0.3 to $0.7\,T_c$. It is most interesting to note that around 5 K I_0 becomes larger than the measured I_c. This could point towards an explanation for the fast deterioration of the detector performance observed at around this temperature, see above. However, at this point it is not clear, if the extrapolated current I_0 has a physical meaning and if the observed temperature dependence can be interpreted as an indication for vortices assisting in the photon detection. Regardless of the physical interpretation, the observation of a temperature-independent γ is an important experimental fact that currently defies a theoretical explanation, see also section 5.4.

4.4. Position dependence of the detection threshold

There have been experimental indications that the detection threshold is position-dependent [15], i.e. the threshold current $I_{\rm th}$ for a given photon energy depends on the absorption position along the cross-section of the nanowire. The existence of such a position-dependent detection efficiency was very recently verified by polarization measurements [99]. Light polarized perpendicularly to the wire is absorbed preferentially in the center of the wire, whereas light polarized parallel to the wire is absorbed uniformly. This enables one the use of differential polarization measurements to compare the detection probability at different parts of the wire. Since the length scale associated with the preferential absorption is itself wavelength dependent, a combination of polarization and wavelength scans allows for a reconstruction of the detection probability with a resolution of 10–20 nm.

Figure 9 shows the result of this reconstruction, as well as a theoretical curve based on the numerical model that will be discussed in more detail in section 5.4. The figure shows the threshold current required to locally obtain 100% detection probability given a photon absorption at that point, for photons with a wavelength $\lambda = 1500$ nm. The experiments revealed a position-dependent threshold current that results in a position-dependent detection efficiency as long as I_b is less than the maximum $I_{\rm th}(y)$, as illustrated in the inset of figure 9. The threshold current near the edges turns out to be $\approx 10\%$ lower than in the center of the wire. This is a very important result and a crucial test for the detection models.

4.5. Magnetic-field dependence of photon counts

The importance of checking the effect of external magnetic fields on the detection of photons in SNSPDs has been realized long ago together with the introduction of the vortex scenario for the detection process [100] and for the dark counts [60]. However, experiments with magnetic field had not been put forward until recently.

First measurements of dark counts in magnetic fields [61, 67, 77] showed that the quasi-static London approach [66] qualitatively describes the observations, see section 3.

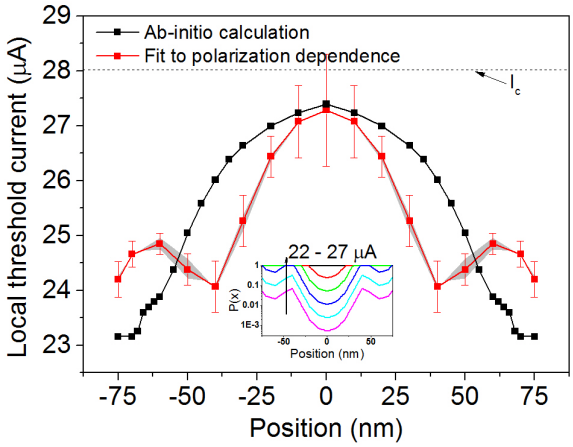

Figure 9. Local threshold current for a 150 nm wide nanodetector, for light with $\lambda = 1500$ nm. The red curve shows the experimental value based on the measurements presented in [99]. The grey band around the data (most prominent around $w = \pm 50$ nm) represents the systematic error induced by the uncertainty in determining the dielectric constant of NbN. The black curve shows the prediction based on the numerical model discussed in section 5.4. (Adapted, with permission, from [99]. © 2015, American Chemical Society.)

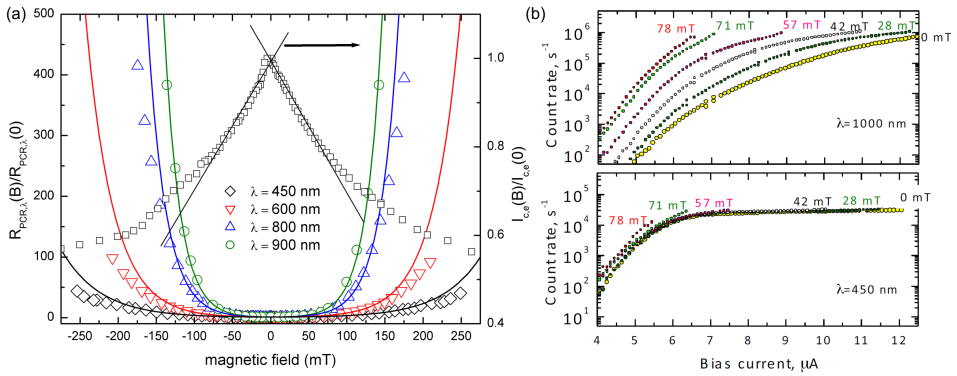

Figure 10. (a) Rates of photon counts for different photon energies and the experimental critical current of a meander in magnetic field (Reproduced with permission from [67]. © 2014, American Physical Society.). (b) Current dependence of photon-count rates for two photon energies and different magnetic fields (© 2015, IEEE. Reprinted, with permission, from [101].). In (a) and (b), one can see that a magnetic field has a stronger effect on the detection of low-energy photons.

In figure 10(a) typical results of photon counts as a function of magnetic field together with the critical current are shown [67]. It is worth noting that the measurements were done on a meander which consists of straight lines and bends. The local current densities in these sections are influenced differently by the external magnetic field as discussed in section 2. While in the Meissner state, the critical current of the meander linearly decreases with the field. The dependence changes the slope and acquires steps when vortices appear in the structure. Qualitatively similar to dark counts, photon-count

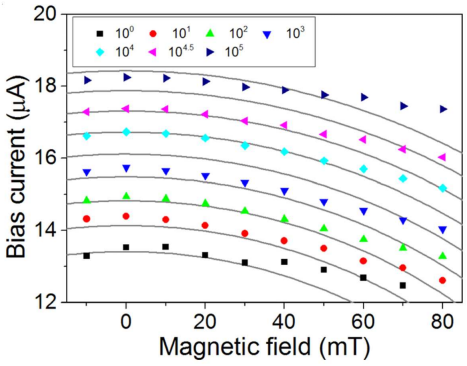

Figure 11. Bias current at constant photon-count rates as a function of the magnetic field (Reproduced with permission from [102]. © 2015, AIP Publishing LLC.). Solid lines describe equidistant ellipses as expected from the Usadel equation. Up to ≈ 50 mT the ellipses can describe the data points. This points at the importance of the variation of the order parameter with applied field and current for the detection of photons.

rates increase with increasing magnetic field.

One can see that the rate of photon counts for high-energy photons is almost field independent in the Meissner state, while the rates for low energy photons start to grow noticeably at much smaller fields. Similar results were obtained with meander structures made from films of MoSi [101], see figure 10(b). There is a strong field-dependence of photon counts for photons with a wavelengths of 1000 nm and almost no change with magnetic field for a wavelength of 450 nm.

Recent experiments with two-dimensional nanobridges in a magnetic field [102] have shown that pairs of the magnetic field and current, which keep the photon-count rates constant, fall onto an elliptic curve in the B–I coordinates at small magnetic fields $\lesssim 50$ mT, see figure 11(a). The authors interpreted the interplay between field and current at small fields in terms of a depairing energy Γ that was inferred from the solution of the 1D-Usadel equation [103]

$$\Gamma \propto \alpha I_b^2 + \beta B^2, \tag{6}$$

with known constants α and β. One can interpret (6) such that current and magnetic field result in an equivalent reduction of the order parameter. Physically, this implies that the order parameter should control the photon-count rates. At magnetic fields above ≈ 50 mT the current density across the bridge becomes non-uniform enough to possibly cause deviations from the 1D-Usadel approach indicated by a systematic deviation from the elliptical curves.

A different behaviour of photon and dark counts, which has been observed in many experiments, is also seen in their magnetic-field dependence. It is most plausibly explained in meanders by different origins for these events. Current assisted photon counts come mostly from straight segments of the meander while fluctuation assisted photon and dark counts originate from bends [29, 61, 67]. This does not exclude the possibility that dark counts in the straight fragments of meanders or in nanobridges

have a different mechanism as compared to photon counts [102].

5. Detection models and comparison with experiments

5.1. Normal-core hotspot model

The first model to describe the detection mechanism of SNSPDs already contains several of the ingredients also used in later models [1, 88, 104]. Absorption of the photon results in an excited electron with high energy as compared to the superconducting gap, $E_{\mathrm{exc}} \approx h\nu \gg \Delta$, with h the Planck constant and ν the photon frequency. This excited electron thermalizes in a cascading process generating a large number of elementary excitations, so-called quasi-particles (QPs) and mostly athermal phonons. The QPs are subject to diffusion resulting in a cloud of QPs with growing diameter. A given density of QPs will locally suppress superconductivity and a normal-conducting core may form in the center of this QP-cloud. Originally, it was assumed that the number of QPs $n_{\mathrm{qp}} \approx h\nu/\Delta$ after the thermalization time τ_{th}. In this case, simulations suggested a normal-core diameter larger than the superconducting coherence length ξ_{GL} of NbN even for near-infrared photons with energies of about 1 eV.

Once such a normal-core with radius $r_{nc} > \xi_{\mathrm{GL}}/2$ has formed, the bias current will be redistributed[+]. The current avoids the normal-core and the current density is assumed to increase uniformly in the still superconducting parts of the cross-section containing the absorption site. If the diameter of the normal-core is big enough, the critical current density will be exceeded and a detection event will occur through the suppression of superconductivity across the wire.

This model is essentially based on an area argument: each unit of energy contributes to making the normal-core larger, which serves to increase the diameter of the obstacle that the current must overcome. Since the system is 2D, the size scales as $\sqrt{E}$, and the energy-current relation has a square-root behaviour:

$$I_{\mathrm{th}} = I_c \left(1 - \gamma' \sqrt{E}\right), \tag{7}$$

where E is the energy of the incident photon, $I_c = I_{c,\mathrm{dep}}$ is the critical current assumed to be the depairing critical current and γ' is the device and material dependent proportionality factor.

It has soon been realized that this model results in quantitative inconsistencies when applied to experimental data [74]. However, the strongest argument against it comes from the observed linear energy-current relation (5), which is incompatible with (7).

[+] Normal-cores with smaller diameter will not result in significant current redistribution, because the superconducting order parameter cannot change significantly on length scales smaller than ξ_{GL}

5.2. Diffusion-based hotspot model

A more sophisticated model was put forward in 2005 [74], that made one important generalization as compared to the normal-core hotspot model: the need for a normal-conducting core of the hotspot is given up. Instead, the number of Cooper pairs in a minimal section of the wire with a length equal to the coherence length, called a ξ-slab, is considered. These Cooper pairs, which are reduced in number compared to the unperturbed superconductor, must still carry the current which was carried by the original number of pairs. Therefore, they must speed up. If the Cooper pairs exceed the critical velocity v_c, they break up and the wire transitions to the normal state. Because the current carrying capacity of the wire is proportional to the number of remaining Cooper-pairs, the energy-current relation is of the form:

$$I_{\mathrm{th}} = I_{c,\mathrm{dep}} \left(1 - \frac{E}{E_0}\right). \tag{8}$$

Applying several simplifications and approximations it is possible to derive an expression for the energy scale [74]

$$E_0 = \gamma'^{-1} = \left(\frac{N(0)\Delta^2 wd}{\varsigma}\right) \sqrt{\pi D \tau_{\mathrm{th}}}, \tag{9}$$

where the dimensionless parameter ς represents the efficiency with which a photon is converted from an initial excitation in the material to QPs at the superconducting-gap edge. It captures, for example, losses to the phonon bath. Since the value of this parameter could in principle differ from film to film, it serves essentially as a fit parameter for each set of experimental observations.

This model (also called the *refined hotspot model* in literature) is in agreement with several experimental observations. First of all, (8) is in accordance with the linear energy-current relation of (5). Equation (9) is also in accordance with the observed linear dependence of the threshold on the width w and thickness d of the nanowire [23, 78], as well as giving a reasonable description of the dependence on the material parameters $N(0)$, Δ and D [84]. And taking into account that not only Δ, but also the diffusion coefficient D of QPs is temperature-dependent, this model can give a qualitatively correct description of the temperature-dependence of the threshold a low temperatures $T/T_c \lesssim 0.5$ [97].

However, this model also predicts a down-shift of the threshold for $T/T_c > 0.5$, which is not observed experimentally. More importantly, according to (8) the relevant current scale should be the depairing critical current, whereas in experiments it is consistently found that $I_0/I_c \approx 0.75 \ldots 0.8$ [78, 94, 95]. Furthermore, the recently confirmed position-dependence of the threshold cannot be explained by this model. A further weakness of both the normal-core hotspot model and the diffusion-based hotspot model is that they predict a deterministic, threshold-like response: the detector responds to all photons of a particular energy, or it does not. In contrast, experimentally it is observed that there is a slow roll-off in the count rate when the bias current through the device is decreased before at even lower bias currents the count-rate decreases

exponentially. Particularly the exponential behaviour points to a regime in which some sort of fluctuation, either thermal or quantum mechanical in nature, occasionally assists in producing a detection event. A natural candidate for these fluctuations are magnetic vortices, which are taken into account in the following detection models.

5.3. Photon-triggered vortex-entry model

In 2008, the notion was put forward that vortex-antivortex pairs (VAP) could be responsible for the slow roll-off of the detection efficiency at longer wavelengths and lower currents than the threshold [100]. Later in 2010, single-vortex hopping was proposed as the detection mechanism beyond the threshold [23], and in 2011, it was suggested that vortices are responsible for all detection events, including those at high currents in the plateau of the detection efficiency [63]. In this model, the arrival of the photon decreases the entry barrier for vortices as discussed in section 3, which enables a vortex crossing or makes it at least more likely. In a follow up paper [66] a complete description of the photon-triggered vortex-entry scenario was worked out, giving analytical expressions for the current, temperature and magnetic-field dependence of the count rates. To do so, the authors used the same London-model approach as for the description of vortex-induced dark counts, see section 3, including those limitations discussed above. Additionally, they made a very simplistic assumption for the effect of the photon on the vortex-entry condition, namely that the photon results in a homogeneous reduction of the order parameter inside a so-called hot belt, but the photon energy is insufficient to make it normal-conducting. Therefore, the bias current is still uniform across the strip. The reduction of the order parameter reduces the energy barrier for vortex entry and for high enough photon energy E at a given I_b it will even vanish completely. Thus, this model gives an explanation for the detection plateau, as well as a more gradual reduction of the detection efficiency below the threshold.

In this model, the functional dependence of the detection probability R in the regime $R \ll 1$ is of the form [66]:

$$R \propto I^{\nu_h+1}, \tag{10}$$

where ν_h is a parameter that measures the reduction of the energy barrier for vortex entry to the absorption of a photon. Following [66,67,78], we apply the assumption that the energy is divided equally over the area of interest. This results in a value of ν_h of:

$$\nu_h = \nu - 4\pi\varsigma E/(k_b T)(\xi/w)^2, \tag{11}$$

where $\nu = \varepsilon_0/(k_b T)$ is the value of the vortex energy scale in the absence of photon absorption in units of the thermal energy $k_B T$. The two expressions above imply an approximate energy-current relation of the form

$$I/I_{c,v} = \exp(C/(\nu - E/E_0 + 1)), \tag{12}$$

with C being a constant, and $I_{c,v}$ is the critical current that reduces the energy barrier for vortex entry to zero. Bear in mind, that the energy scale E_0 in (12) is different from (8).

Within the model of [66] $I_{c,v} \approx 0.8 I_{c,\mathrm{dep}}$, thus it may explain the observed $I_0 \approx 0.75 I_{c,\mathrm{dep}}$, if one associates I_0 with $I_{c,v}$. However, it has to be noted that the value of $I_{c,v}$ depends on the vortex model and vortex-core energy used to calculate the critical current for a vanishing vortex-entry barrier. On the other hand, the highly nonlinear relation (12) is clearly incompatible with the experimentally determined linear energy-current relation. Furthermore, from (11) one can estimate the dependence of the minimum photon energy E_{min} on the wire width $E_{\mathrm{min}} \propto w^2$, which is also not supported by experiments [78].

This model also allows one to derive analytical approximations for the magnetic-field dependence of dark and photon counts. At magnetic fields much smaller than a characteristic field value the energy barrier decreases approximately linear with the applied magnetic field. This results in an approximately exponential increase in the rate of vortex entry from one edge. Also taking into account anti-vortices that enter from the opposite edge one expects the rate [66]

$$R(H) \propto \cosh \frac{H}{H_1}, \tag{13}$$

with different field scales H_1 for dark and photon counts, respectively, from which one can also derive the vortex-barrier energy scales. Qualitatively, this behaviour is confirmed in experiments, see for example figure 10(a), where the photon-count rates are fitted to cosh-relations. Whereas the vortex-barrier energy scale for dark counts as determined from least-square fits to the experimental field-dependence is reasonably close to the theoretical expectation [61], it turns out that the experimental energy scale for photon counts is about one order of magnitude different compared to theoretical values [67, 101, 102]. This may be a consequence of the various simplifications used in this theoretical model. Nevertheless, the fitted value of the vortex-barrier energy scale for photon counts varied in the expected way with the photon energy and the current [67].

5.4. Diffusion-based vortex-entry model

The diffusion-based hot-spot model (section 5.2) could successfully describe many experimental observations. Its weak point is the assumption that the local velocity of the superconducting condensate is uniform across the strip. A model was proposed [105], with some later corrections [106], that combines the generation and diffusion of QPs after photon absorption with the formation of the normal-conducting cross-section due to vortices overcoming the edge-barrier for vortex entry. It turns out that a realistic description of the underlying processes can only be obtained with numerical methods. Because of that some previously applied simplifications could be given up as well.

The model assumes that the whole process can be separated into three independent processes. It starts with the QP generation and diffusion. Instead of using the analytical solutions for extended 2D-films, the diffusion equation is solved numerically on an infinitely long strip. Furthermore, the model distinguishes between the excited electron after photon absorption and thermalized QPs. The excited electron is assumed to diffuse

with a temperature-independent diffusion coefficient equal to that of electrons in the normal state D_e, and it is assumed that this excited electron generates QPs with a constant time-scale τ_{qp} resulting in an exponential increase in the number of QPs. The diffusion of QPs on the other hand is governed by a temperature-dependent coefficient D_{qp} and QPs will eventually recombine to form Cooper-pairs on a time-scale τ_r. The phonon temperature is assumed to stay constant on the relevant short time scale of τ_{th} that is assumed to be the time at which the maximum total number of QPs are present in the film. The background of thermally excited QPs is neglected. This results in the coupled differential equations [105]

$$\frac{\partial C_e(\vec{r},t)}{\partial t} = D_e \nabla^2 C_e(\vec{r},t) \tag{14}$$

$$\frac{\partial C_{qp}(\vec{r},t)}{\partial t} = D_{qp}\nabla^2 C_{qp}(\vec{r},t) - \frac{C_{qp}(\vec{r},t)}{\tau_r} + \frac{\varsigma h\nu}{\Delta \tau_{qp}} \exp\left(-\frac{t}{\tau_{qp}}\right) C_e(\vec{r},t), \tag{15}$$

with ∇^2 being the Laplace-operator, $C_e(\vec{r},t)$ is the probability density to find the excited electron at position $\vec{r}$ at time t, and $C_{qp}(\vec{r},t)$ is the local density of excess QPs at time t. This simple exponential behaviour for the increase in the number of QPs was used instead of a more detailed description [88, 104, 107] to avoid unnecessary complexity. The inhomogeneous distribution of QPs may also be interpreted as a continuously varying order parameter of the superconducting condensate. The order parameter varies comparatively smoothly in the strip even after photon absorption, because the typical size of the QP-cloud, which may be expressed by the diffusion length $L_d = \sqrt{Dt}$ with an effective coefficient $D_{qp} \leq D \leq D_e$, is larger than the coherence length ξ_{GL} almost immediately after photon absorption.

Similar to the diffusion-based hot-spot model, the bias current redistribution is approximated to be instantaneous, due to the relevant time scales being much shorter than τ_{qp}. Therefore, the current-density distribution is estimated by solving the Laplace equation [106]

$$\vec{\nabla} \cdot \left(n_{se} \vec{\nabla} \varphi\right) = 0, \tag{16}$$

where φ is the phase of the superconducting order parameter connected to the drift velocity of superconducting electrons $\vec{v}_s = (\hbar/m)\vec{\nabla}\varphi$, m is assumed to be the free electron mass. The current density is then $\vec{j} = n_{se}\vec{v}_s$, where it is assumed that the local sum of the density of superconducting electrons n_{se} and the local density of QPs C_{qp} equals the equilibrium density of superconducting electrons. The smooth variation of n_{se} and of the order parameter, respectively, result in a significantly reduced peak of the current density due to the current-crowding effect * as compared to the expected current-crowding around a hole in a superconducting strip [28].

The inhomogeneous order parameter and current distribution, respectively, also necessitate a numerical calculation of the vortex potential. This is done by placing a test-vortex at different positions across the superconducting strip [106]. As in the

* See appendix in [99]

previous vortex model the London-model for a vortex is used neglecting the contribution of the vortex core to the vortex self-energy. Following [28] the vortex self-energy is determined by calculating the total current circulating the vortex-core and the work done by the external current is obtained by integrating the bias-current density from the point of entry to the current position of the test-vortex. For a given photon energy and absorption position the detection threshold is determined as the bias-current density $I_b/I_{c,\mathrm{dep}}$ that reduces the edge-barrier to zero.

In [105] numerical results were compared with implementations of the normal-core hotspot model and the diffusion-based hotspot model within the same framework. The authors found clear indications that their new model and the underlying physics would lead to the entry of a vortex and thus the trigger of a normal domain before either of the other models predicts a detection event. Most importantly, numerical results for the energy-current relation are fully compliant with the linear energy-current relation (5) obtained with QDT. As it turns out, the numerical model and the algorithms used to solve the differential equations start to produce systematic errors for photon energies that are too high. Therefore, the calculations are restricted to photon energies smaller than about 1 eV corresponding to photon wavelengths $\gtrapprox$ 1200 nm. These photon energies are still high enough to achieve some overlap with experimentally obtained data for which the lowest photon energies are typically $\sim$ 0.8 eV and extrapolations using (5) allow a direct comparison between numerical and experimental results. Good agreement between theory and experiment is usually found by variations of the energy conversion efficiency ς and small adjustments to the zero-energy extrapolated threshold current I_0. These adjustments of I_0 are within the estimated numerical and experimental accuracies, respectively.

A possible position-dependence of the threshold criterion has been investigated using this numerical model [106]. In figure 12(a) the calculated threshold currents I_{th} as a function of the photon absorption position across a 150 nm wide NbN strip are shown for a selection of photon wavelengths. The data are symmetric around the center line due to the symmetry between vortices and anti-vortices. The decrease in I_{th} for absorption positions near the edge is significant. At each absorption position the energy-current relation is linear to a very good approximation. As can be seen in figure 12(b) the magnitude of the slope is increasing towards the edge, whereas the extrapolated value for $E = 0$ remains constant within the uncertainties. This is highlighted in the inset of figure 12(b). The averaged I_0 is almost identical with the vortex-entry current in the absence of a photon, indicated by the red and black horizontal lines in figure 12(a), respectively. This position-dependence of the threshold gives a direct explanation for the rounding observed in the universal detection curve [94] and the current-dependence of the detection efficiency, respectively. As one increases the bias current starting from low values, at first the areas near the edge of the superconducting strips will reach the maximum detection efficiency. One has to increase I_b further until the threshold current for the center to obtain maximum detection efficiency for the complete strip. In the inset of figure 12(a) a comparison is shown between detection probabilities obtained in

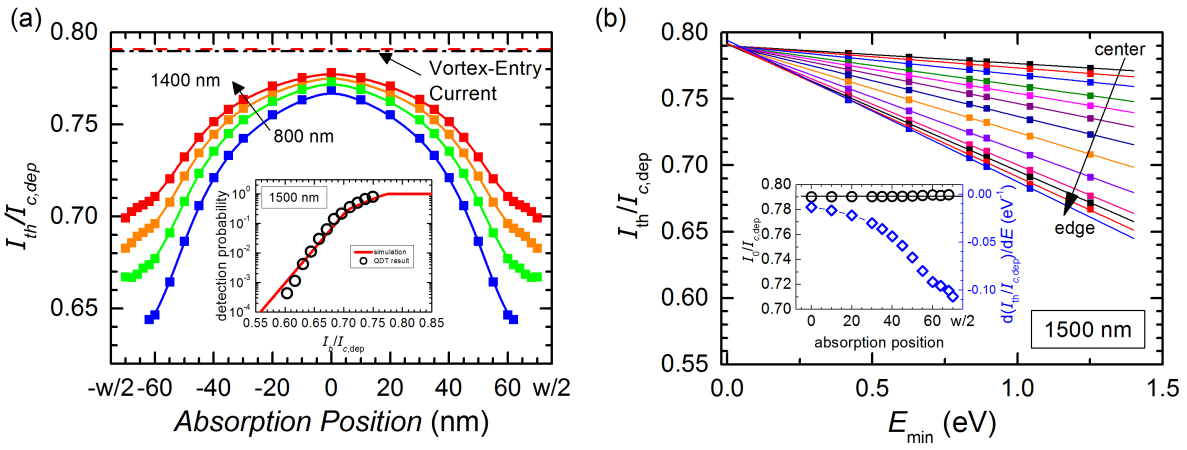

Figure 12. (a) Calculated dependence of $I_{\rm th}$ on the photon absorption position in a long, 150 nm-wide NbN strip for several photon wavelengths (data from [99]). The black dash-dotted line indicates the vortex-entry current obtained from the calculations. *Inset*: detection probability for 1500 nm-photons in a 150 nm wide NbN strip determined with QDT (black circles) compared with the prediction from the numerical model based on the position-dependence of $I_{\rm th}$ (red line). (b) The linear energy-current relation extracted from the data in (a) for each absorption position. The vortex-entry current I_0 in the limit of zero photon-energy is to a good approximation independent of the absorption position. This is highlighted in the inset, where I_0 (black data, left axis) and the slope of the linear relation (blue data, right axis) are plotted versus the position. The average of I_0 (red dashed line in the inset of (a)) coincides with the vortex-entry current in the absence of a photon as calculated directly from the simulations within numerical uncertainties.

a QDT experiment with the expectation from the model calculations as detailed in the appendix of [99].

As mentioned in section 4.2, there have been indications of a lower threshold for photon detection for absorption events near the edge, which could recently been verified in polarization-dependent QDT experiments [99]. In figure 9 the calculated position-dependent threshold current $I_{\rm th}$ for photons of 1500 nm wavelength absorbed in a 150 nm wide NbN strip is plotted together with the reconstructed position-dependence from the experiment. The model gives a fair description of the overall experimental observations. The model reproduces correctly that the threshold current is maximal in the center of the strip and is reduced towards the edge by about 10%. Nevertheless, there are some important differences that may shed some light on weak points of the current detection model. The model simulations result in a near monotonic decrease all the way to edge, whereas experimental data suggest that $I_{\rm th}$ levels off for distances less than $\approx$30 nm from the edge and a narrower peak at the center. At the current stage the model assumes *perfect* edges. The edges of real devices are far from perfect and could be a reason for the observed discrepancies.

Further indications come from simulations of 2-photon events that the model does not capture all effects that are of importance for photon absorptions near the edge [106]. QDT experiments have proven that the detector response is determined only by the

total absorbed energy irrespective of the number of photons. In order to check that this important result can be reproduced simulations were run with two 2000 nm-photons absorbed at random positions and the threshold-current distribution has been compared with threshold currents obtained for 1-photon absorptions of 1000 nm wavelength. At high bias-currents when the detection probability exceeds ≈ 0.5 the two calculated detection probabilities or universal curves practically coincide. However, at lower bias-currents or detection probabilities there are systematic deviations. At these lower bias currents detection events come primarily from absorption events close to the edge. The fact that in this regime the simulations result in systematic differences between 1- and 2-photon events indicate that the current model is insufficient to capture all effects that are important at the edge.

There are some other experimental observations, for which the simulation results produce the general trend, but fail to give the correct functional relations. The model correctly results in an approximately linear up-shift of the threshold for increasing width w, but cannot quantitatively reproduce the $E_{\min} \propto w$ dependence observed experimentally [78]. Similarly, the T-dependence of the threshold as computed from the model qualitatively correctly predicts a down-shift of the threshold for decreasing T below $\approx 0.5T_c$ [97], but it cannot explain the T-dependence of I_0 deduced from QDT [95]. In fact, the numerical results would indicate a T-independent I_0 and a position- and T-dependent slope γ, in contradiction to the experimental observations, see also section 4.3.

5.5. Normal-core vortex model

Using an alternative approach, a detection model has been developed [108] that sees a role for vortices entering from the edge as well as VAPs being generated inside the superconducting strip. In their original proposal they assume a circular or semicircular area in the center or at the edge of the strip, respectively, with a fixed radius, where the effective temperature is increased due to the absorption of the photon. For this situation they solve the time-dependent Ginzburg-Landau (TDGL) equation, the heat-diffusion equation and Poisson's equation for the electrical potential to determine the evolution of the superconducting state after photon absorption and the formation of the voltage pulse that signals the detection event. However, this first approach resulted in several predictions that are in contradiction with experimental observations, e.g. a non-linear energy-current relation, an extrapolated zero-energy threshold current $I_0 = I_{c,\mathrm{dep}}$, $E_{\min} \propto w^2$, and a lower threshold current for photon absorption in the center than at the edge of the strip. It is well known that the temperature range for which the TDGL is strictly applicable is much more restricted to $T \approx T_c$ than for the GL equation [109], and it could be a reason that this approach failed to describe several experimental results.

A more recent model [110, 111] can describe the detection in an SNSPD more successfully. It is still assumed that photon absorption results in a well-defined area with an increased effective temperature. This area is of circular or semicircular shape depending on the absorption position towards the center or very close to the edge,

respectively. In order to simplify their model, they neglect the evolution in time of the hotspot and use a fixed radius for the circular area

$$\varsigma E \simeq d\pi R^2 \frac{H_{cm}^2}{8\pi}, \tag{17}$$

with H_{cm} the thermodynamic critical magnetic field and $H_{cm}^2/8\pi$ the superconducting condensation energy per unit volume. The semicircles at the edge will have a $\sqrt{2}$ times larger radius to keep the volume constant. Inside these hotspots the excess QPs are modelled by a nonequilibrium distribution function using a single parameter $\alpha = const < 1$ inside the hotspot and $\alpha = 1$ outside. The order parameter will vary smoothly inside and outside of the hotspot, due to the superconducting proximity effect. They then use a relaxation method to find stationary solutions of the GL equation. This has the advantage that they do not need to make any assumption about the mechanism resulting in a resistive cross-section, instead it can be deduced from the evolution of the order parameter.

The results obtained within this model can be interpreted in the following way [111]: for photon absorptions close to the edge, there is a minimum current I_{en} that allows the entry of a single vortex from the edge into the hotspot. If the hotspot radius $R \gtrsim 3\,\xi_{\mathrm{GL}}$ a slightly larger current I_{pass} is required to overcome the pinning force the vortex experiences inside the hotspot. Only then is the vortex free and can move across the strip and cause a detection event. This situation prevails until the gap between the hotspot and the near edge of the strip exceeds $\sim 2\,\xi_{\mathrm{GL}}$. Is the photon absorbed further away from the edge a slightly different scenario is suggested. Then, there is a critical current I_{pair} above which a VAP is formed inside the hotspot. Once more, a slightly larger current I_{pass} is required to overcome the binding energy of the VAP. This latter scenario can be easily motivated by considering a strip of $w/2$ and a semicircular notch defect at one edge as done in [28]. Current-crowding around the defect favours the entry of a vortex at the point of minimum cross-section. The case of a hot-spot in the center of a strip of width w with strong suppression of the order parameter may then be seen as two mirrored strips with a notch, joined together.

Using this revised model Vodolazov calculated the threshold current I_{th} for a detection probability ≈ 1 as a function of photon energy [110]. For the energy range probed by QDT [94, 95] the model results in a nearly linear energy-current relation compatible with experimental data, see also figure 13(a). However, the model predicts significant deviations from the linear energy relation at photon energies $E \lesssim 0.5$ eV and a zero-energy extrapolated $I_0 = I_{c,\mathrm{dep}}$. This deviation from linearity for $E \to 0$ is claimed to be a consequence of a realistic description of the vortex core in the GL-model as compared to the London-model, but has so far not been observed in experiments ♯.

The position-dependence of I_{th} has also been calculated in the normal-core vortex

♯ The data presented in [112] for a NbN detector are in contradiction to all data obtained with QDT. In the same publication, the energy-current relation determined for WSi obtained from data published in [113] contains errors and can also not been seen as an indication for a nonlinear energy-current relation.

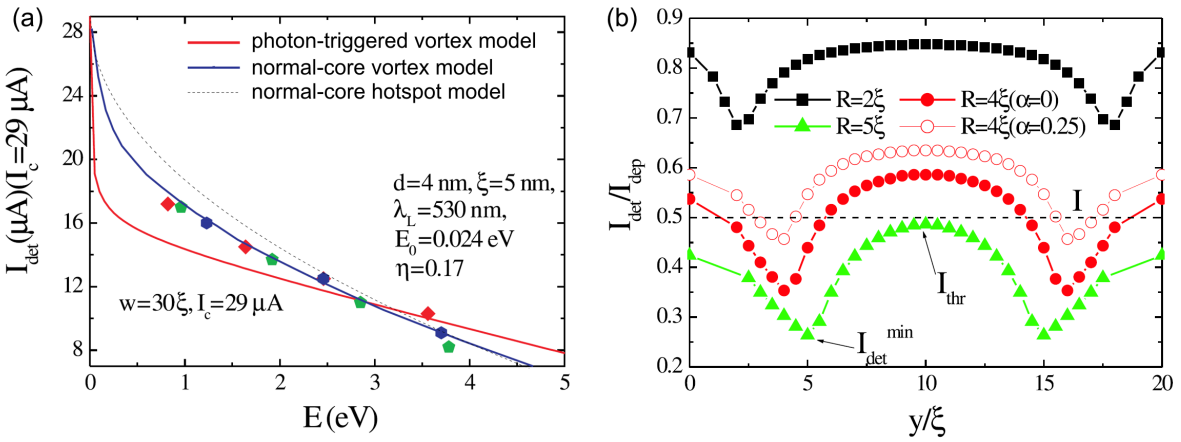

Figure 13. (a) Least-square fit of the energy-current relation of the normal-core vortex model to experimental data (taken from [94]) in comparison with two other models as indicated (Adapted, with permission, from [110]. © 2014, American Physical Society.). (b) Position-dependence of the threshold current ($I_{\rm det}$ in the figure) for different hotspots. The strip width $w = 20\,\xi_{\rm GL}$ is roughly equivalent to 100 nm (Reproduced with permission from [111]. © IOP Publishing. All rights reserved.).

model [111], the results are shown in figure 13(b). Calculated $I_{\rm th}$ have a broad maximum near the center of the strip and $I_{\rm th}$ decreases sharply towards the edge with a minimum for distances equal to the radius R of the hotspot. For absorption positions even closer to the edge $I_{\rm th}$ increases again up to values almost as high as in the center. It could be that the minimum at a distance R from the edge is an artifact of the calculations, due to the abrupt change from a circular hotspot to a hotspot in the shape of a truncated circle. From these position-dependent threshold currents detection probabilities can be deduced which are similar compared to those computed from the diffusion-based vortex model and experimentally measured detection probabilities, see figures 12 and 6, respectively.

However, compared to the experimentally derived position-dependence shown in figure 9, the calculated threshold currents within the normal-core vortex model deviate significantly. The dome around the center is flatter than experimentally observed for an even wider strip. The differences between this model and the measurements of the position-dependence are more significant for absorption positions closer to the edge. The minimum $I_{\rm th}$ at a distance R from the edge is 20% or more reduced compared to the center and the sharp increase very close to the edge up to approximately the same value of $I_{\rm th}$ as in the center has not been seen in experiments. The last revision of the model [112] results in a some quantitative changes of the position-dependence. Although for some model parameters the minimum $I_{\rm th}$ is now reduced by an amount similar to the experimental results, the threshold current very close the edges is predicted to exceed $I_{\rm th}$ in the center by more than 20%.

On the other hand, numerical results published in [110] are consistent with the linear relation between $E_{\rm min}$ and w determined experimentally [78] in a range of bias

currents that correspond to about 0.4 to $0.6 I_{c,\mathrm{dep}}$. At higher or lower bias currents the simulated results suggest a nonlinear relation between E_{min} and w. Unfortunately, for these bias current ranges, no experimental data are available.

The normal-core vortex model can also qualitatively explain experimental magnetic-field data. It predicts that the strongest effect of the magnetic field on the rate of photon counts should occur at detection probabilities < 1, where the light counts are due to fluctuation assisted or current assisted vortex entry. Interestingly, this model predicts a slight decrease in the photon-count rate at small magnetic fields and bias currents just below I_{th}, figure 11 in [111], a consequence of the position-dependence of I_{th} on the position of the hot-spot. It may explain the small increase of the photon-count rate seen in figure 3 of [67] near zero magnetic field. In a very recent preprint [112] the model has been further adapted and compared with new experimental data. The decreasing photon-count rate with increasing magnetic field at bias currents just below I_{th} could be confirmed theoretically and experimentally. The model can now also reproduce the very weak magnetic field effect on count rates of high energy UV photons, see lower panel of figure 10(b).

The magnetic field-dependence of photon counts measured at low fields on nanobridges [102] could be well explained by the 1D-Usadel equations. At least for the 1D-case it was shown that the GL and the Usadel approaches result in very similar depairing factors at temperatures above $0.5\,T_c$ [114]. It is therefore possible that these two approaches may turn out to be complimentary to each other even for the 2D-case of typical SNSPDs.

6. Conclusions

Above discussions have demonstrated that there is overwhelming experimental and theoretical evidence for a significant role of magnetic vortices in the detection of photons and the generation of dark counts in SNSPDs. Additionally, the success of the diffusion-based hotspot model, e.g. in the description of the width-dependence and the linear energy-current relation, is a strong indication of the importance of QP diffusion early on in the detection process. In many situations this model may be sufficient to give guidance on how to improve a detector for a certain application. With (9) it is possible to qualitatively predict the effective change of a device or material parameter on the detector threshold. However, this model cannot explain the observed position-dependence of the detection threshold. The same can be said about the photon-triggered vortex-entry model, which can be regarded as an extension of the diffusion-based model to treat the entry of vortices and the effects of magnetic fields in an analytical way. The latter one can successfully predict the qualitative functional dependence of the photon and dark counts on an applied magnetic field, but it fails to quantitatively describe the photon-count rates.

There are currently two models requiring numerical solutions of the underlying equations that allow a more detailed description of the photon detection process in

SNSPDs. The diffusion-based vortex-entry model applies a phenomenological approach. This model predicts a linear energy-current relation that is in excellent agreement with the results from QDT. It can also reasonably well reproduce the position-dependent threshold current as determined from polarization dependent experiments. On the other hand, there are also experimental results which cannot be explained yet, e.g. the width and temperature dependencies. There are indications that the model produces systematic errors when the photon is absorbed too close to the edge. Currently, the boundary conditions imply perfect edges, a condition that real structures cannot fulfill. Structural defects, changes in stoichiometry or oxidation are all likely to result in normal-conducting areas at the edges of the strip [115]. As a consequence of the superconducting proximity effect, there will be edge regions with a reduced order parameter that act as QP traps and reduce the effective number of QPs. One may speculate that the model can be improved by taking into account these edge effects, but then one should also consider QP trapping inside the photon generated QP cloud. The use of the London-approximation for the vortex-core could be another reason that the diffusion-based vortex-model cannot explain all experimental observations.

The normal-core vortex-model is the most fundamental of all proposed detection models. Based on the argument that QPs are trapped inside the hotspot due to a rapidly reduced order parameter, an area with a constant, non-equilibrium distribution function for the QPs is assumed and the Ginzburg-Landau equation is solved for this situation, either in its time-independent or its time-dependent form. The model does not use the simple London approximation of vortices and it also does not need to make any assumption about the mechanism that leads to the formation of the normal domain to trigger a detection event. When the photon is absorbed sufficiently close to the edge the solutions suggest the entry of a single vortex from the edge just like the vortex models that use the London approximation. It differs from the latter quantitatively, however. For photon absorptions near the center of the strip a slightly different mechanism is proposed, in which a VAP is generated inside the hotspot. If the bias current is strong enough the pair unbinds and both vortices move towards opposite edges and trigger a normal domain. This model can reproduce the energy-current relation obtained from QDT reasonably accurate and the newest implementation can very successfully explain the magnetic field dependence of photon counts. However, the proposed position-dependence of the threshold current differs systematically from the results derived with the polarization experiment. It seems likely that the particular position-dependence of this model is an artefact of the change of the hotspot shape from circular to semicircular for distances less than the hotspot radius R from the edge.

In table 1 we attempt to give a quick overview by comparing the different detection models with experimental observations. The normal-core hot-spot model is not included in this list, since it cannot describe the detector response beyond a basic phenomenological description.

From the experimental point of view the most interesting challenge is to measure the energy-current relation at lower photon energies towards the mid-infrared. The

normal-core vortex model predicts a pronounced deviation from the currently observed linear behaviour contrary to the diffusion-based model. Alternatively, one may compare a range of superconducting materials for which the deviation from the linear energy-current relation may occur at higher photon energies. Most likely this would be the case for materials for which the detection threshold is comparatively high.

SNSPDs have also been used to detect particles of much higher energies [116, 117] with clear indications that a simple normal-core hotspot model is sufficient to explain the detection mechanism. An interesting question here is the cross-over from a vortex-based detection mechanism to a mechanism that does not involve vortices. The recently proposed amorphous WSi [113] is particularly interesting in this respect, because the low T_c and low density of states $N(0)$ suggest the formation of a large normal-conducting area upon absorption of relatively low-energy photons in the near-infrared [106]. Furthermore, the extended, very flat plateau in the detection efficiency and the possibility to operate WSi-SNSPD up to $0.7T_c$ are outstanding differences of this material compared all other materials suggested so far.

Another interesting question is the time delay between photon absorption and the formation of the normal-conducting domain that triggers the detection event. The diffusion-based as well as the normal-core vortex model can in principle give an answer to this question. However, for both models a more accurate description of the QP multiplication process and the formation of the hotspot, respectively, will be required to obtain reliable results. With the current simulations one may estimate it to be $\lesssim \tau_{\rm th}$, which is as short as a few picoseconds. Ultimately, statistical or position-dependent variations of this time delay may put a limit on the smallest achievable timing jitter of SNSPDs. Exploiting correlation effects in 2-photon effects first attempts have been made to determine the size and life-time of the QP-cloud and hotspot, respectively [93]††. There are also a number of open questions concerning the dark counts in SNSPD, already raised above during the discussion of dark counts. It is therefore highly likely that SNSPDs will continue to stimulate both experimentalists and theoreticians to investigate metastable and nonequilibrium states in these superconducting structures.

Acknowledgments

J.J.R. acknowledges the NWO (Spinoza award). All authors acknowledge stimulating discussions with D. Vodolazov.

††Just before the final submission of this paper, experimental and theoretical work has been published on the size and life-time of the hot-spot in WSi detectors [118, 119].

Table 1. Comparison of the various detection models with experimental observations. The first row states whether the model has analytical or numeric solutions. A checkmark (✓) means that the model can describe the experimental observation reasonable well, a cross (x) means the model fails to reproduce experimental data and a dash (—) means this property has not been or cannot be analyzed in this model.

exp. observation	diffusion hot-spot model	simple vortex-entry model	diffusion vortex-entry model	normal-core vortex model
	analytic	analytic	numeric	numeric
linear $E_{\min}$–I_{th}	✓[a]	x	✓	✓[b]
position-dependence	x	x	✓	x
linear w–$E_{\min}$	✓	x	(✓)[c]	✓[b]
T-dependence	(✓)[d]	—	x[e]	—
H-dependence	—	(✓)[f]	—	✓
universal curve	x	x	✓	✓

[a] at $E = 0$ model predicts $I_{\mathrm{th}}(0) = I_{c,\mathrm{dep}}$

[b] model predicts nonlinear behaviour outside experimentally measured range

[c] linearity is reproduced, but quantitatively wrong

[d] reasonable description of the down-shift of the threshold at low T

[e] predicts a down-shift of the threshold at low T

[f] describes functional dependence, but quantitatively wrong

Appendix A. Quantum detector tomography

Appendix A.1. Introduction

In this appendix, we give a brief review ‡ of quantum detector tomography [89, 90]. The goal of performing a QDT experiment is to find the response of the detector in the number state (Fock) basis, i.e. to find out the probability of a detection event p_n if the detector were to receive exactly n photons as an input.

Unfortunately, the vast majority of optical sources produce states which have some distribution in the photon number basis, i.e. which are not eigenstates of the photon number operator. It is therefore not possible to directly measure p_n. Instead, a statistical approach is adopted, in which the detector is illuminated with a series of states with known distributions in the photon number basis §. A numerical transformation is then applied to convert these results into the response in the number state basis.

The most convenient set of states to use for quantum detector tomography is the set of coherent states. These states, which are produced by a laser [120], have the convenient property that an attenuated coherent state still remains coherent. This means that it

‡ The presentation of the material follows [93]

§ Since the detector is not senstive to the phase of the incoming photon, we may adopt a classical picture in which we consider only the photon number probability distribution, i.e. we restrict ourselves to the diagonal elements of the density matrix.

is possible to create the desired set of photon number probability distributions simply by attenuation ‖.

The key idea of tomography is that it is independent of any preconceived notions of the working mechanism of the detector. The quantities p_n which are obtained from tomography are reproducible numbers that do not depend on any model of the detection mechanism. This makes this approach particularly suited for the investigation of the detection mechanism of SNSPDs.

Appendix A.2. Experimental considerations

A typical tomography experiment consists of a pulsed laser which is fed to a variable attenuator. This light is then fed to the device under test. Count rates are recorded for each light intensity. In the case of SNSPDs, typically, the bias current through the device is also swept, but each bias current constitutes a separate experiment.

The main experimental demands of QDT are on the light source. In particular, the pulses which the light source sends out should be shorter than the lifetime of an excitation in the detector. This number can be determined in a separate experiment [121] and is 20 ps for NbN SNSPDs. Moreover, the pulses should be separated by more than the dead time of the detector, to make sure that each individual pulse corresponds to a separate probing of the detector, uncorrelated to the previous ones. It should also be understood that the detection probabilities obtained from tomography apply to the mode with the specific temporal, spatial and frequency characteristics that was used to perform tomography.

For fiber-based experiments, part of the power is tapped off via a beam splitter to a power monitor, which tracks fluctuations in the laser intensity. In this way, slow power fluctuations, i.e. those on a time scale comparable to the experiment duration, can be monitored. Feito *et al.* [90] have considered the effect of pulse-to-pulse fluctuations and found that these can by accounted for by a broadening of the photon number distribution.

The demands on the attenuator are as follows: experimentally, we find that between 3 and 4 orders of magnitude of attenuation are sufficient to resolve the quantities p_n of interest, i.e. those unequal to zero or one [99]. For large attenuations, which are typically necessary for highly efficient detectors, the strategy is to use several attenuators and calibrate each attenuator separately. This avoids two problems: first, the fact that bolometric power meters have a dark current which defines a minimum level of intensity which can be measured, typically 20 pW (corresponding to 10^8 photons/s), and second: the fact that power meters are typically nonlinear when measuring over a large dynamic range.

‖ Alternative sources have been proposed, but these have found little application beyond the single-photon level

Appendix A.3. Mathematics of QDT

In the most general case, the detector has multiple detection outcomes. An example of this is a multi-element detector where one or more of the elements of the device may fire, producing a voltage pulse whose height is proportional to the number of elements which fired. Indexing the outcomes by m and the number of photons which caused them by n, and noting that in general we offer up a series of light states indexed i, we may write:

$$\mathbf{R} = \mathbf{CP}, \tag{A.1}$$

where $\mathbf{R} = R_{m,i}$ is the probability of the m-th kind of outcome to the i-th input state, $\mathbf{C} = c_{n,i}$ is the matrix that contains the probability that state i contains n photons, and $\mathbf{P} = p_{n,m}$ is the probability that n photons will result in the m-th kind of detection event. Since we are using coherent states, which are fully characterized by their mean photon number N and have $c_{n,i} = \exp(-N_i)N_i^n/n!$, we may write:

$$R_m(N_i) = \sum_n e^{-N_i} \frac{(N_i)^n}{n!} p_{n,m} \tag{A.2}$$

Appendix A.4. Numerical inversion

A general issue with tomography problems is that they are numerically unstable, i.e. the matrix $\mathbf{C}$ is generally close to singular. This means that in general, a least-squares fit $\mathbf{P} = \mathbf{C}^{-1}\mathbf{R}$, where $\mathbf{C}^{-1}$ is the Penrose-Moore pseudoinverse, will strongly amplify small errors $\Delta\mathbf{R}$ onto the error $\Delta\mathbf{P}$. There are several approaches in literature to this problem. All of these approaches involve adding some weak constraints to the problem to make the inversion more tractable.

The earliest approach, which was developed by Lundeen, *et al.* [89]. is to add the assumption that the p_n are smooth in n, i.e. that if we know how the detector responds to n photons, we can make a decent guess how it will respond to $n+1$ photons: $p_{n+1} \approx p_n$. This constraint is enforced by penalizing the solution for large differences between adjacent p_n by adding an extra term $k\sum_n (p_n - p_{n+1})^2$ to the least squares fit. This has the effect of penalizing solutions which have large difference between photon numbers, which are considered unphysical in this approach.

A second approach is to introduce the weak assumption about the detector that some part of the action of the detector is linear loss [91]. This linear loss corresponds to a reduction of the mean photon number. We can therefore write:

$$R_m(N_i) = \sum_n e^{-\eta N} \frac{(\eta N_i)^n}{n!} p_{n,m}, \tag{A.3}$$

where $0 \leq \eta \leq 1$ is the parameter describing the linear loss. Equation (A.3) is now overdetermined: only products of η and p_n enter. To resolve this issue, an additional assumption is introduced, which is sparsity: the assumption that only few p_n will have values unequal to 0 or 1. This approach can be formalized in the form of model selection,

which leads to a rigorous definition of the notion of 'the model that fits the data well without too many parameters'. In this way, a linear efficiency η and a set of detection probabilities p_n can be obtained simultaneously. This approach was found to be particularly useful for detectors with small overall efficiency or small active area.

References

[1] Gol'tsman G N, Okunev O, Chulkova G, Lipatov A, Semenov A, Smirnov K, Voronov B, Dzardanov A, Williams C and Sobolewski R 2001 *Appl. Phys. Lett.* **79** 705

[2] Eisaman M D, Fan J, Migdall A and Polyakov S V 2011 *Rev. Sci. Instrum.* **82** 071101

[3] Natarajan C M, Tanner M G and Hadfield R H 2012 *Supercon. Sci. Technol.* **25** 063001

[4] Dauler E A, Grein M E, Kerman A J, Marsili F, Miki S, Nam S W, Shaw M D, Terai H, Verma V B and Yamashita T 2014 *Opt. Eng.* **53** 081907

[5] Engel A, Bartolf H, Gómez L, Semenov A, Il'in K, Schilling A, Hübers H W and Siegel M 2009 *THE THIRTEENTH INTERNATIONAL WORKSHOP ON LOW TEMPERATURE DETECTORS LTD13* (*AIP Conf. Proc.* vol 1185) ed Young B, Cabrera B and Miller A pp 76–79

[6] Marsili F, Najafi F, Dauler E, Bellei F, Hu X, Csete M, Molnar R J and Berggren K K 2011 *Nano Letters* **11** 2048–2053

[7] Halperin B I and Nelson D R 1979 *J. Low Temp. Phys.* **36** 599

[8] Halperin B I, Refael G and Demler E 2010 *Intern. J. Mod. Phys. B* **24** 4039–4080

[9] Pearl J 1964 *Appl. Phys. Lett.* **5** 65

[10] Maksimova G M 1998 *Phys. Solid State* **40** 1607

[11] Clem J R 1998 *Bull. Am. Phys. Soc.* vol 43 p 411

[12] Stan G, Field S B and Martinis J M 2004 *Phys. Rev. Lett.* **92** 097003

[13] Benkraouda M and Clem J R 1998 *Phys. Rev. B* **58** 15103

[14] Jones W A, Barnes P N, Mullins M J, Baca F J, Emergo R L S, Wu J, Haugan T J and Clem J R 2010 *Appl. Phys. Lett.* **97** 262503

[15] Anant V, Kerman A J, Dauler E A, Yang J K W, Rosfjord K M and Berggren K K 2008 *Opt. Express* **16** 10750

[16] Driessen E F C, Braakman F R, Reiger E M, Dorenbos S N, Zwiller V and de Dood M J A 2009 *Eur. Phys. J. Appl. Phys.* **47** 10701

[17] Baek B, Stern J A and Nam S W 2009 *Appl. Phys. Lett.* **95** 191110

[18] Miki S, Takeda M, Fujiwara M, Sasaki M and Wang Z 2009 *Opt. Express* **17** 23557–23564

[19] Liu D, Miki S, Yamashita T, You L, Wang Z and Terai H 2014 *Opt. Express* **22** 21167–21174

[20] Kerman A J, Dauler E A, Keicher W E, Yang J K W, Berggren K K, Gol'tsman G and Voronov B 2006 *Appl. Phys. Lett.* **88** 111116

[21] Yang J K W, Kerman A J, Dauler E A, Anant V, Rosfjord K M and Berggren K K 2007 *IEEE Trans. Appl. Supercon.* **17** 581 –585

[22] Haas P, Semenov A, Hübers H W, Beyer J, Kirste A, Schurig T, Il'in K, Siegel M, Engel A and Smirnov A 2007 *IEEE Trans. Appl. Supercon.* **17** 298

[23] Hofherr M, Rall D, Il'in K, Siegel M, Semenov A, Hübers H W and Gippius N A 2010 *J. Appl. Phys.* **108** 014507

[24] You L, Yang X, He Y, Zhang W, Liu D, Zhang W, Zhang L, Zhang L, Liu X, Chen Sand Wang Z and X i X 2013 *AIP Advances* **3** 072135

[25] Henrich D, Dörner S, Hofherr M, Il'in K, Semenov A, Heintze E, Scheffler M, Dressel M and Siegel M 2012 *J. Appl. Phys.* **112** 074511

[26] Gershenzon M E and Gubankov V N 1980 *JETP Lett.* **32** 217

[27] Neumann L G and Kao Y H 1982 *J. Low Temp. Phys.* **48** 321

[28] Clem J R and Berggren K K 2011 *Phys. Rev. B* **84**(17) 174510

[29] Akhlaghi M K, Atikian H, Eftekharian A, Loncar M and Majedi A H 2012 *Opt. Express* **20** 23610–23616
[30] Henrich D, Reichensperger P, Hofherr M, Meckbach J M, Il'in K, Siegel M, Semenov A, Zotova A and Vodolazov D Y 2012 *Phys. Rev. B* **86**(14) 144504
[31] Hortensius H L, Driessen E F C, Klapwijk T M, Berggren K K and Clem J R 2012 *Appl. Phys. Lett.* **100** 182602
[32] Il'in K, Henrich D, Luck Y, Liang Y, Siegel M and Vodolazov D Y 2014 *Phys. Rev. B* **89** 184511
[33] Semenov A, Günther B, Böttger U, Hübers H W, Bartolf H, Engel A, Schilling A, Ilin K, Siegel M, Schneider R, Gerthsen D and Gippius N A 2009 *Phys. Rev. B* **80** 054510
[34] Yang J K W, Kerman A J, Dauler E A, Cord B, Anant V, Molnar R J and Berggren K K 2009 *IEEE Trans. Appl. Supercon.* **19** 318–322
[35] Yamashita T, Miki S, Terai H and Wang Z 2013 *Opt. Express* **21** 27177
[36] Clem J R, Mawatari Y, Berdiyorov G R and Peeters F M 2012 *Phys. Rev. B* **85**(14) 144511
[37] Maksimova G M, Zhelezina N V and Maksimov I L 2001 *Europhys. Lett.* **53** 639–645
[38] Elistratov A A, Vodolazov D Y, Maksimov I L and Clem J R 2002 *Phys. Rev. B* **66** 220506
[39] Adami O A, Cerbu D, Cabosart D, Motta M, Cuppens J, Ortiz W A, Moshchalkov V V, Hackens B, Delamare R, Van de Vondel J and Silhanek A V 2013 *Appl. Phys. Lett.* **102** 052603
[40] Sobolewski R, Verevkin A, Gol'tsman G N, Lipatov A and Wilsher K 2003 *IEEE Trans. Appl. Supercon.* **13** 1151
[41] Korneev A, Kouminov P, Matvienko V, Chulkova G, Smirnov K, Voronov B, Goltsman G N, Currie M, Lo W, Wilsher K, Zhang J, Słysz W, Pearlman A, Verevkin A and Sobolewski R 2004 *Appl. Phys. Lett.* **84** 5338
[42] Engel A, Semenov A, Hübers H W, Ilin K and Siegel M 2004 *Nucl. Instrum. Methods A* **A520** 32
[43] Semenov A, Engel A, Il'in K, Siegel M and Hübers H W 2005 *IEEE Trans. Appl. Supercon.* **15** 518
[44] Kitaygorsky J, Zhang J, Verevkin A, Sergeev A, Korneev A, Matvienko V, Kouminov P, Smirnov K, Voronov B, Gol'tsman G and Sobolewski R 2005 *IEEE Trans. Appl. Supercon.* **15** 545
[45] Bell M, Sergeev A, Mitin V, Bird J, Verevkin A and Gol'tsman G 2007 *Phys. Rev. B* **76** 094521
[46] Engel A, Semenov A, Hübers H W, Il'in K and Siegel M 2005 *phys. stat. sol. (c)* **2** 1668
[47] Engel A, Semenov A D, Hübers H W, Il'in K and Siegel M 2006 *Phys. C* **444** 12
[48] Kitaygorsky J, Komissarov I, Jukna A, Pan D, Minaeva O, Kaurova N, Divochiy A, Korneev A, Tarkhov M, Voronov B, Milostnaya I, Gol'tsman G and Sobolewski R R 2007 *IEEE Trans. Appl. Supercon.* **17** 275–278
[49] Langer J S and Ambegaokar V 1967 *Phys. Rev.* **164** 498
[50] McCumber D E and Halperin B I 1970 *Phys. Rev. B* **1** 1054
[51] Skocpol W J and Tinkham M 1975 *Rep. Prog. Phys.* **38** 1049
[52] Little W A 1967 *Phys. Rev.* **156** 396–403
[53] Likharev K K 1979 *Rev. Mod. Phys.* **51** 101
[54] Berezinskii Z L 1970 *Zh. Eksp. Teor. Fiz. [Sov. Phys. JETP]* **59** 907
[55] Berezinskii Z L 1971 *Zh. Eksp. Teor. Fiz. [Sov. Phys. JETP]* **61** 1144
[56] Kosterlitz J M and Thouless D J 1973 *J. Phys. C* **6** 1181
[57] Mooij J E 1984 *Percolation, Localization, and Superconductivity* ed Goldman A M and Wolf S A (Plenum Press New York) p 325
[58] Benkraouda M and Clem J R 1996 *Phys. Rev. B* **53** 5716
[59] Maksimov I L and Maksimova G M 1997 *JETP Lett.* **65** 423
[60] Bartolf H, Engel A, Schilling A, Il'in K, Siegel M, Hübers H W and Semenov A 2010 *Phys. Rev. B* **81** 024502
[61] Engel A, Schilling A, Il'in K and Siegel M 2012 *Phys. Rev. B* **86** 140506(R)
[62] Qiu C and Qian T 2008 *Phys. Rev. B* **77** 174517
[63] Bulaevskii L N, Graf M J, Batista C D and Kogan V G 2011 *Phys. Rev. B* **83** 144526

[64] Gurevich A and Vinokur V M 2012 *Phys. Rev. B* **86**(2) 026501
[65] Vodolazov D Y 2012 *Phys. Rev. B* **85**(17) 174507
[66] Bulaevskii L N, Graf M J and Kogan V G 2012 *Phys. Rev. B* **85**(1) 014505
[67] Lusche R, Semenov A, Korneeva Y, Trifonov A, Korneev A, Gol'tsman G and Hübers H W 2014 *Phys. Rev. B* **89** 104513
[68] Nasti U, Parlato L, Ejrnaes M, Cristiano R, Taino T, Myoren H, Sobolewski R and Pepe G 2015 *Phys. Rev. B* **92**(1) 014501
[69] Yamashita T, Miki S, Makise K, Qiu W, Terai H, Fujiwara M, Sasaki M and Wang Z 2011 *Appl. Phys. Lett.* **99** 161105
[70] Murphy A, Semenov A, Korneev A, Korneeva Y, Gol'tsman G and Bezryadin A 2015 *Sci. Rep.* **5** 10174
[71] Tafuri F, Kirtley J R, Born D, Stornaiuolo D, Medaglia P G, Orgiani P, Balestrino G and Kogan V G 2006 *Europhys. Lett.* **73** 948
[72] Hofherr M, Rall D, Il'in K, Semenov A, Hübers H W and Siegel M 2012 *J. Low Temp. Phys.* **167**(5) 822–826
[73] Verevkin A, Zhang Z, Sobolewski R, Lipatov A, Okunev O, Chulkova G, Korneev A, Smirnov K, Gol'tsman G N and Semenov A 2002 *Appl. Phys. Lett.* **80** 4687
[74] Semenov A, Engel A, Hübers H W, Il'in K and Siegel M 2005 *Eur. Phys. J. B* **47** 495
[75] Gaudio R, op 't Hoog K P M, Zhou Z, D S and Fiore A 2014 *Appl. Phys. Lett.* **105** 222602
[76] Noat Y, Cherkez V, Brun C, Cren T, Carbillet C, Debontridder F, Il'in K, Siegel M, Semenov A, Hübers H W and Roditchev D 2013 *Phys. Rev. B* **88** 014503
[77] Lusche R, Semenov A, Il'in K, Korneeva Y, Trifonov A, Korneev A, Hübers H W, Siegel M and Gol'tsman G 2013 *IEEE Trans. Appl. Supercond.* **23** 2200205
[78] Lusche R, Semenov A, Il'in K, Siegel M, Korneeva Y, Trifonov A, Korneev A, Gol'tsman G, Vodolazov D and Hübers H W 2014 *Journal of Applied Physics* **116** 043906
[79] Zhang L, Kang L, Chen J, Zhong Y, Zhao Q, Jia T, Cao C, Jin B, Xu W, Sun G and Wu P 2010 *Applied Physics B* **102** 867–871
[80] Maingault L, Tarkhov M, Florya I, Semenov A, Espiau de Lamaëstre R, Cavalier P, Gol'tsman G, Poizat J P and Villégier J C 2010 *J. Appl. Phys.* **107** 116103
[81] Marsili F, Bitauld D, Fiore A, Gaggero A, Mattioli F, Leoni R, Benkahoul M and Lévy F 2008 *Opt. Express* **16** 3191–3196
[82] Miki S, Takeda M, Fujiwara M, Sasaki M, Otomo A and Wang Z 2009 *Applied Physics Express* **2** 075002
[83] Dorenbos S N, Forn-Díaz P, Fuse T, Verbruggen A H, Zijlstra T, Klapwijk T M and Zwiller V 2011 *Appl. Phys. Lett.* **98** 251102
[84] Engel A, Aeschbacher A, Inderbitzin K, Schilling A, Il'in K, Hofherr M, Siegel M, Semenov A and Hübers H W 2012 *Appl. Phys. Lett.* **100** 062601
[85] Marsili F, Verma V B, Stern J A, Harrington S, Lita A E, Gerrits T, Vayshenker I, Baek B, Shaw M D, Mirin R P and Nam S W 2013 *Nature Photonics* **7** 210–214
[86] Korneeva Y P, Mikhailov M Y, Pershin Y P, Manova N N, Divochiy A V, Vakhtomin Y B, Korneev A A, Smirnov K V, Devizenko A Y and Goltsman G N 2014 *Supercond. Sci. Technol.* **27** 095012
[87] Verma V B, Lita A E, Vissers M R, Marsili F, Pappas D P, Mirin R P and Nam S W 2014 *Appl. Phys. Lett.* **105** 022602
[88] Semenov A D, Gol'tsman G N and Korneev A A 2001 *Phys. C* **351** 349
[89] Lundeen J S, Feito A, Coldenstrodt-Ronge H, Pregnell K L, Silberhorn C, Ralph T C, Eisert J, Plenio M B and Walmsley I A 2008 *Nat. Phys.* **5** 27–30
[90] Feito A, Lundeen J S, Coldenstrodt-Ronge H, Eisert J, Plenio M B and Walmsley I A 2009 *New J. Phys.* **11** 093038
[91] Renema J J, Frucci G, Zhou Z, Mattioli F, Gaggero A, Leoni R, de Dood M J A, Fiore A and van Exter M P 2012 *Opt. Express* **20** 2806–2813

[92] Bitauld D, Marsili F, Gaggero A, Mattioli F, Leoni R, Jahanmirinejad S, Lévy F and Fiore A 2010 *Nano Lett.* **10** 2977–81
[93] Renema J J 2015 *The physics of nanowire superconducting single-photon detectors* Ph.D. thesis Department of Quantum Matter & Optics, Leiden Institute of Physics (LION), Faculty of Science, Leiden University
[94] Renema J J, Frucci G, Zhou Z, Mattioli F, Gaggero A, Leoni R, de Dood M J A, Fiore A and van Exter M P 2013 *Phys. Rev. B* **87** 174526
[95] Renema J J, Gaudio R, Wang Q, Zhou Z, Gaggero A, Mattioli F, Leoni R, Sahin D, de Dood M J A, Fiore A and van Exter M P 2014 *Phys. Rev. Lett.* **112** 117604
[96] Yamashita T, Miki S, Qiu W, Fujiwara M, Sasaki M and Wang Z 2010 *Applied Physics Express* **3** 102502
[97] Engel A, Inderbitzin K, Schilling A, Lusche R, Semenov A, Hübers H W, Henrich D, Hofherr M, Il'in K and Siegel M 2013 *IEEE Trans. Appl. Supercon.* **23** 2300505
[98] Verma V B, Korzh B, Bussires F, Horansky R D, Lita A E, Marsili F, Shaw M D, Zbinden H, Mirin R P and Nam S W 2014 *Appl. Phys. Lett.* **105** 122601
[99] Renema J, Wang Q, Gaudio R, Komen I, op 't Hoog K, Sahin D, Schilling A, van Exter M, Fiore A, Engel A and de Dood M 2015 *Nano Letters* **15** 4541–4545
[100] Semenov A D, Haas P, Günther B, Hübers H W, Ilin K and Siegel M 2008 *J. Low Temp. Phys.* **151** 564
[101] Korneev A A, Korneeva Y P, Mikhailov M Y, Pershin Y P, Semenov A V, Vodolazov D Y, Divochiy A, Vakhtomin Y B, Smirnov K V, Sivakov A, Devizenko A Y and Gol'tsman G N 2015 *IEEE Trans. Appl. Supercon.* **25** 2200504
[102] Renema J J, Rengelink R J, Komen I, Wang Q, Gaudio R, op 't Hoog K P M, Zhou Z, Sahin D, Fiore A, Kes P, Aarts J, van Exter M P, de Dood M J A and Driessen E F C 2015 *Appl. Phys. Lett.* **106** 092602
[103] Anthore A, Pothier H and Esteve D 2003 *Phys. Rev. Lett.* **90** 127001
[104] Semenov A D, Gol'tsman G N and Sobolewski R 2002 *Supercon. Sci. Technol.* **15** R1
[105] Engel A and Schilling A 2013 *J. Appl. Phys.* **114** 214501
[106] Engel A, Lonsky J, Zhang X and Schilling A 2015 *IEEE Trans. Appl. Supercon.* **25** 2200407
[107] Kozorezov A G, Volkov A F, Wigmore J K, Peacock A, Poelaert A and den Hartog R 2000 *Phys. Rev. B* **61** 11807
[108] Zotova A N and Vodolazov D Y 2012 *Phys. Rev. B* **85**(2) 024509
[109] Kopnin N 2001 *Theory of Nonequilibrium Superconductivity* (*International Series of monographs in physics* vol 110) (Oxford University Press)
[110] Vodolazov D Y 2014 *Phys. Rev. B* **90** 054515
[111] Zotova A N and Vodolazov D Y 2014 *Supercond. Sci. Technol.* **27** 125001
[112] Vodolazov D Y, Korneeva Y P, Semenov A V, Korneev A A and Goltsman G N 2015 Vortex assisted mechanism of photon counting in superconducting nanowire single photon detector revealed by external magnetic field arXiv:1505.04689
[113] Baek B, Lita A E, Verma V and Nam S W 2011 *Appl. Phys. Lett.* **98** 251105
[114] Semenov A V, Krutitskii P A and Devyatov I A 2010 *JETP Letters* **92**(11) 762–766
[115] Il'in K S, Siegel M, Semenov A D, Engel A and Hübers H W 2004 *Applied Superconductivity 2003, Proc. of the 6th Eur. Conf. on Appl. Supercond. Sorrento, Italy* (*Institute of Physics Conference Series* no 181) ed Andreone A, Pepe G P, Cristiano R and Masullo G p 2895
[116] Suzuki K, Shiki S, Ukibe M, Koike M, Miki S, Wang Z and Ohkubo M 2011 *Applied Physics Express* **4** 083101
[117] Inderbitzin K, Engel A and Schilling A 2013 *IEEE Trans. Appl. Supercond.* **23** 2200505
[118] Marsili F, Stevens M J, Kozorezov A, Verma V B, Lambert C, Stern J A, Horansky R, Dyer S, Duff S, Pappas D P, Lita A, Shaw M D, Mirin R P and Nam S W 2015 Hotspot Relaxation Dynamics in a Current Carrying Superconductor arXiv:1506.03129
[119] Kozorezov A G, Lambert C, Marsili F, Stevens M J, Verma V B, Stern J A, Horansky R, Dyer

S, Duff S, Pappas D P, Lita A, Shaw M D, Mirin R P and Nam S W 2015 *accepted for publ. in PRB*

[120] Glauber R 1963 *Phys. Rev.* **131** 2766

[121] Zhou Z, Frucci G, Mattioli F, Gaggero A, Jahanmirinejad S, Hoang T B and Fiore A 2013 *Phys. Rev. Lett.* **110** 133605